

```

Query Match      20.5%; Score 61.5; DB 12; Length 131;
Best Local Similarity 35.3%; Pred. No. 0.89;
Matches 18; Conservative 7; Mismatches 11; Indels 15; Gaps 4;

QY      3 NHKGVAFRH-CN-----PNGT-----WDFMHSLNKTKWANSDCLEFLQPDIS 43
DB      29 NSMGCAFRGVCHTVEWPNGTLPQNW------NNTWQEWKRYVRELEANT 74

```

Search completed: November 8, 2000, 08:56:04
Job time: 515 sec

Db 38 NSWGAFRQVCHTTPVWTFNSTPKWDNMTWQEWEDQVRFLFANIS 82

RESULT 11

Q06268 ID Q06268 PRELIMINARY; PRT; 591 AA.
AC Q06268;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE INTERMEDIATE FILAMENT PROTEIN.
OS Octopus dofleini (Giant Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
OC Incirrata; Octopodidae; Octopus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LENS;
RX MEDLINE; 94060097.
RA Tomarev S.I., Zinovleva R.D., Platigorsky J.;
RT "Primary structure and lens-specific expression of genes for an
intermediate filament protein and a beta-tubulin in cephalopods.";
RL Biochim. Biophys. Acta 1216:245-254(1993).
CC -1- TISSUE SPECIFICITY: LENS.
DR EMBL; LI0113; AAA16613.1; -.
DR INTERPRO; IPR000531; -.
DR INTERPRO; IPR001322; -.
DR INTERPRO; IPR001664; -.
DR PFAM; PF00038; filament; 2.
DR PFAM; PF00932; IF_tail; 1.
DR PROSITE; PS00226; IF; 1.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
KW Intermediate filament; Eye lens protein.
SQ SEQUENCE 591 AA; 67287 MW; 4A72245E20C6F57E CRC64;

Query Match 22.3%; Score 67; DB 5; Length 591;

Best Local Similarity 35.9%; Pred. No. 0.83;

Matches 14; Conservative 8; Mismatches 17; Indels 0; Gaps 0;

Qy 3 NHKGVAFRHCNPNGTWDFMHSLNKWTWYSDCLRFLQPD 41

Db 52 SHKGVTDIRCNRKREKKEKLQDLNRFANYIEKVRFLAE 90

RESULT 12

Q9WUP2 ID Q9WUP2 PRELIMINARY; PRT; 463 AA.
AC Q9WUP2;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CALCITONIN RECEPTOR-LIKE RECEPTOR.
GN CRUR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA Derst C., Preisig-Mueller R., Gerhardus J., Daut J.;
RT "Cloning and sequencing of mouse CGRP/adrenomedullin receptor
subunits."
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF146525; AAD35021.1; -.
DR INTERPRO; IPR000832; -.
DR INTERPRO; IPR001688; -.
DR INTERPRO; IPR001879; -.
DR PFAM; PF00002; 7tm.2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PRINTS; PR00361; CALCITONINR.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 463 AA; 53234 MW; 630EC8956A58847E CRC64;

Query Match 20.7%; Score 62; DB 11; Length 463;

Best Local Similarity 37.1%; Pred. No. 2.9;

Matches 13; Conservative 6; Mismatches 14; Indels 2; Gaps 2;

Qy 1 DFNHKGVAFRHCNPNGTWDFMH-SLNKWTWYSDC 34

Db 93 DFDPSEKVTIKCDQDQGHW-FRHPDSNRTWTNTYTL 126

RESULT 13

Q9RIW5 ID Q9RIW5 PRELIMINARY; PRT; 463 AA.
AC Q9RIW5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CALCITONIN RECEPTOR-LIKE RECEPTOR PRECURSOR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SPLEEN;
RA Tsujikawa K., Tadotsu N., Takizawa A., Hayashi T., Yamamoto H.;
RT "Mouse calcitonin receptor-like receptor."
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB015595; BAA76492.1; -.
DR INTERPRO; IPR000832; -.
DR INTERPRO; IPR001688; -.
DR INTERPRO; IPR001879; -.
DR PFAM; PF00002; 7tm.2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PRINTS; PR00361; CALCITONINR.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 463 AA; 53204 MW; 338CC0B5ED2B6899 CRC64;

Query Match 20.7%; Score 62; DB 11; Length 463;

Best Local Similarity 37.1%; Pred. No. 2.9;

Matches 13; Conservative 6; Mismatches 14; Indels 2; Gaps 2;

Qy 1 DFNHKGVAFRHCNPNGTWDFMH-SLNKWTWYSDC 34

Db 93 DFDPSEKVTIKCDQDQGHW-FRHPDSNRTWTNTYTL 126

RESULT 14

Q9QXH8 ID Q9QXH8 PRELIMINARY; PRT; 463 AA.
AC Q9QXH8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CALCITONIN RECEPTOR-LIKE RECEPTOR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57 BL/6; TISSUE-LUNG;
RA Ono Y., Okano I., Kojima M., Okada K., Kangawa K.;
RT "cDNA cloning of mouse CRLR and RAMPs."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF209905; AAF21037.1; -.
DR INTERPRO; IPR000832; -.
DR INTERPRO; IPR001688; -.
DR INTERPRO; IPR001879; -.
DR PFAM; PF00002; 7tm.2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.


```
DE COMPLEMENT FACTOR C2.
GN C2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=129;
RA Rowen L., Madan A., Qin S., Hall J., Dahl T., James R., Dickhoff R.,
RA Schaffer T., Ratcliffe A., Abbasi N., Loretz C., Lasky S., Hood L.;
RT "Sequence of the mouse MHC class III region.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 291-760 FROM N.A.
RA Rowen L., Qin S., Lasky S.R., Loretz C., Dors M., Mahairas G.,
RA Hood L.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF109906; AAC84162.1; -.
DR EMBL; AF049850; AAC05284.1; -.
DR HSSP; P00734; 2HNT.
DR INTERPRO; IPR000436; -.
DR INTERPRO; IPR001254; -.
DR INTERPRO; IPR001314; -.
DR INTERPRO; IPR002035; -.
DR PFAM; PF00084; sushi; 2.
DR PFAM; PF00089; trypsin; 2.
DR PFAM; PF00092; vwa; 1.
DR PRINTS; PR00453; VWFADOMAIN.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
SQ SEQUENCE 760 AA; 84741 MW; 91C896A3EDC7D448 CRC64;

Query Match                22.8%; Score 68.5; DB 11; Length 760;
Best Local Similarity      33.3%; Pred No. 0.69;
Matches 15; Conservative   7; Mismatches 18; Indels 5; Gap

Qy 1 DFNHKGVAFRCHPNCTGTFMHSLNKTWANYSDCLRFQPDISIG 45
    || :| | | | | | | | | | | | | | | | | | | |
Db 125 DFTLRGSPVYCRPNGLWDGETAVCDNGASHCP-----NPGISVG 164

RESULT 10
O56350 PRELIMINARY; PRT; 150 AA.
AC O56350;
AT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE GP41 (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 2.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PO2;
RA Heredia A.;
RL AIDS Res. Hum. Retroviruses 0:0-0(1998).
DR EMBL; AF039491; AAB99980.1; -.
DR INTERPRO; IPR000328; -.
DR PFAM; PF00517; GP41; 1.
FT NON_TER 1
FT NON_TER 150
SQ SEQUENCE 150 AA; 17976 MW; FDAL177DA2A092EAA CRC64;

Query Match                22.3%; Score 67; DB 12; Length 150;
Best Local Similarity      37.8%; Pred. No. 0.2;
Matches 17; Conservative   6; Mismatches 18; Indels 4; Gap

Qy 3 NHKGVAFRH-CNPNGTWDFMHSU----NKTWANYSDCLEFLQPDIS 43
    | | | | | | | | | | | | | | | | | | | | | |
```


Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	145	48.3	595	6	Q9TU31		Q9TU31 canis fami
2	143	47.7	536	13	Q9PVD3		Q9PVD3 brachydanio
3	140	46.7	542	13	Q9PWD2		Q9PWD2 brachydanio
4	112	37.3	575	13	Q9PWB7		Q9PWB7 brachydanio
5	86	28.7	561	5	Q9VXR9		Q9VXR9 drosophila
6	75	25.0	444	5	Q9VG65		Q9VG65 drosophila
7	70	23.3	553	4	Q95838		Q95838 homo sapien
8	69	23.0	550	11	Q920W0		Q920W0 ratu norv
9	68.5	22.8	760	11	O70350		O70350 mus musculu
10	67	22.3	150	12	Q56350		Q56350 human immu
11	67	22.3	591	5	Q06268		Q06268 octopus dof
12	62	20.7	463	11	Q9WUP2		Q9WUP2 mus musculu
13	62	20.7	463	11	Q9RLW5		Q9RLW5 mus musculu
14	62	20.7	463	11	Q9QXH8		Q9QXH8 mus musculu
15	61.5	20.5	131	12	Q87643		Q87643 chimpanzee
16	61	20.3	681	12	Q36429		Q36429 marburg vir
17	59.5	19.8	388	5	Q9V6C7		Q9V6C7 drosophila
18	59.5	19.8	2155	12	Q64958		Q64958 avian infec
19	59	19.7	854	12	Q73306		Q73306 human immu

RESULT 2

THIS PAGE BLANK (USPTO)

the type 1 receptor.":

J. Neurochem. 69:1640-1649(1997).

-!- FUNCTION: THIS IS A RECEPTOR FOR CORTICOTROPIN RELEASING FACTOR. SHOWS HIGH-AFFINITY BINDING FOR UROTENSIN I. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL CYCLASE (BY SIMILARITY).

-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

-!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch) or send an email to license@isb-sib.ch).

EMBL; Y14037; CAA74364.1; -

CC GCRDB; GCR_2578; -

DR INTERPRO; IPR000832; -

DR INTERPRO; IPR003051; -

DR INTERPRO; IPR003053; -

DR PRAM; PF000002; 7tm_2; 1.

DR PRINTS; PR00249; GPCRSECRETIN.

DR PRINTS; PR01279; CRFRECEPTOR.

DR PRINTS; PR01281; CRFRECEPTOR2.

DR PROSITE; PS00649; G_PROTEIN_RECP_F2_1; 1.

DR PROSITE; PS00650; G_PROTEIN_RECP_F2_2; 1.

CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.

KW SIGNAL 1 ?

FT CHAIN ? 413 CORTICOTROPIN RELEASING FACTOR RECEPTOR

FT FT

FT DOMAIN ? 120 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 121 141 1 (POTENTIAL).

FT DOMAIN 142 150 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 151 170 2 (POTENTIAL).

FT DOMAIN 171 187 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 188 211 3 (POTENTIAL).

FT DOMAIN 212 225 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 226 247 4 (POTENTIAL).

FT DOMAIN 248 266 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 267 289 5 (POTENTIAL).

FT DOMAIN 290 312 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 313 332 6 (POTENTIAL).

FT DOMAIN 333 347 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 348 367 7 (POTENTIAL).

FT DOMAIN 368 413 CYTOPLASMIC (POTENTIAL).

FT CARBOHYD 16 16 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC SEQUENCE 413 AA; 48458 MW; DAD422F0A96C4626 CRC64;

Query Match 22.5%; Score 67.5; DB 1; Length 413;

Best Local Similarity 35.1%; Pred. No. 0.57;

Matches 13; Conservative 4; Mismatches 13; Indels 7; Gaps 1;

Qy 2 FNHKGVAFRHCNPGTWDFMHSLNKTKWANSYSDCLRFL 38

Db 76 YNTRNRYRECFENGW-----ASWMNYSQCVPI 105

Search completed: November 8, 2000, 09:03:45

Job time: 855 sec

FT	STGNAL	1	22	POTENTIAL.
FT	CHAIN	23	478	CALCITONIN RECEPTOR.
FT	DOMAIN	23	153	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	154	173	1 (POTENTIAL).
FT	DOMAIN	174	180	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	181	200	2 (POTENTIAL).
FT	DOMAIN	201	220	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	221	243	3 (POTENTIAL).
FT	DOMAIN	244	260	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	261	280	4 (POTENTIAL).
FT	DOMAIN	281	296	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	297	320	5 (POTENTIAL).
FT	DOMAIN	321	343	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	344	361	6 (POTENTIAL).
FT	DOMAIN	362	373	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	374	395	7 (POTENTIAL).
FT	DOMAIN	396	478	CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	28	28	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	73	73	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	125	125	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	130	130	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	SEQUENCE	478 AA;	55553 MW;	AF22030A94FE2BB4 CRC64;

Query Match 23.3%; Score 70; DB 1; Length 478;
Best Local Similarity 36.8%; Pred. No. 0.33;
Matches 14; Conservative 8; Mismatches 14; Indels 2; Gaps 2;

QY	1	DFNHKGVAFRHCPNGTWDPMH-SLNTWANYSDCLRF	37
	:	::: :: :	
DB	101	DFDPTKVTKYCDSEGVW-FKHPENNRTWSNYTLCNAF	137

RESULT	12
CGRR_HUMAN	STANDARD; PRT; 461 AA.
ID	CGRR_HUMAN
AC	Q16602;
DT	01-NOV-1997 (Rel. 35, Created)
DT	01-NOV-1997 (Rel. 35, Last sequence update)
DT	01-NOV-1997 (Rel. 35, Last annotation update)
DE	CALCITONIN GENE-RELATED PEPTIDE TYPE 1 RECEPTOR PRECURSOR (CGRP TYPE 1 RECEPTOR).
DE	RECEPTOR.
OS	CGRP.
GN	Os homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=LUNG;
RX	MEDLINE; 96212201.
RA	Aiyar N., Rand K., Elshourbagy N.A., Zeng Z., Adamou J.E.,
RA	Bergsma D.J., Li Y.;
RT	"A cDNA encoding the calcitonin gene-related peptide type 1 receptor.";
RL	J. Biol. Chem. 271:11325-11329(1996).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=CEREBELLUM;
RX	MEDLINE; 95118359.
RA	Fluehmann B., Muff R., Hunziker W., Fischer J.A., Born W.;
RT	"A human orphan calcitonin receptor-like structure.";
RL	Biochem. Biophys. Res. Commun. 206:341-347(1995).
CC	!- FUNCTION: THIS IS A RECEPTOR FOR CALCITONIN GENE-RELATED PEPTIDE TYPE 1. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL CYCLASE.
CC	!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC	!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE LUNG AND HEART.
CC	!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

DR	PRINTS: PR00361; CALCITONINR.	
DR	PROSITE; PS00649; G_PROTEIN_RECP_F2.1; 1.	
DR	PROSITE; PS00650; G_PROTEIN_RECP_F2.2; 1.	
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.	
FT	SIGNAL 1 24 POTENTIAL.	
FT	CHAIN 25 515 CALCITONIN RECEPTOR.	
FT	DOMAIN 25 153 EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM 154 173 1 (POTENTIAL).	
FT	DOMAIN 174 180 CYTOPLASMIC (POTENTIAL).	
FT	TRANSMEM 181 200 2 (POTENTIAL).	
FT	DOMAIN 201 257 EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM 258 280 3 (POTENTIAL).	
FT	DOMAIN 281 297 CYTOPLASMIC (POTENTIAL).	
FT	TRANSMEM 298 317 4 (POTENTIAL).	
FT	DOMAIN 318 333 EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM 334 357 5 (POTENTIAL).	
FT	DOMAIN 358 380 6 (POTENTIAL).	
FT	TRANSMEM 381 398 7 (POTENTIAL).	
FT	DOMAIN 399 410 EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM 411 432 7 (POTENTIAL).	
FT	DOMAIN 433 515 CYTOPLASMIC (POTENTIAL).	
SEQ	SEQUENCE 515 AA: 60167 MW: ACB90D2DC3CF166F CRC64;	
Query Match 25.0%; Score 75; DB 1; Length 515;		
Best Local Similarity 39.5%; Pred. No. 0.09;		
Matches 15; Conservative 7; Mismatches 14; Indels 2;		
QY	1 DFNHGVAFRCHNPNGTWFMH-SLNKTWANYSDCLRF 37	
	: : : : : : : : : : : : :	
DB	101 DFDTAERKSVKYCDENGW-FRHDPDSNRKTSWNTLCNAF 137	
RESULT 10		
CALR_HUMAN	STANDARD;	PRT; 490 AA.
ID	CALR_HUMAN	
AC	P30988; O13941; O14585;	
DT	01-JUL-1993 (Rel. 26, Created)	
DT	01-JUL-1993 (Rel. 26, Last sequence update)	
DT	15-JUL-1999 (Rel. 38, Last annotation update)	
DE	CALCITONIN RECEPTOR PRECURSOR (CT-R).	
GN	CALCR.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.	
[1]	SEQUENCE FROM N.A.	
RC	TISSUE=OVARIAN CARCINOMA;	
RX	MEDLINE; 93055406.	
RA	Gorn A.H., Lin H.Y., Yamin M., Auron P.E., Flannery M.R.	
RA	Tapp D.R., Manning C.A., Lodish H.F., Krane S.M., Goldring S.R.;	
RT	"Cloning, characterization, and expression of a human calcitonin	
RL	receptor from an ovarian carcinoma cell line.";	
RL	J. Clin. Invest. 90:1726-1735(1992).	
[2]	SEQUENCE FROM N.A.	
RP	MEDLINE; 94359487.	
RA	Kuestner R.E., Elrod R.D., Grant F.J., Hagen F.S., Kulijper J.L.;	
RA	Mathews S.L., O'Hara P.J., Sheppard P.O., Stroop S.D., Thompson	
RA	Whitmore T.E., Findlays D.M., Houssamis S., Sexton P.M., Moore	
RT	"Cloning and characterization of an abundant subtype of the human	
RT	calcitonin receptor";	
RL	Mol. Pharmacol. 46:246-255(1994).	
[3]	SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.	
RP	TISSUE=BREAST CARCINOMA;	
RC	MEDLINE; 96079881.	
RX	Albrandt K., Brady E.M.G., Moore C.X., Mull E., Sierzega M.E.;	
RA	Beaumont K.;	
RT	"Molecular cloning and functional expression of a third isoform	
RT	of the human calcitonin receptor and partial characterization of the	
RT	calcitonin receptor gene.";	
RL	Endocrinology 136:5377-5384(1995).	

PFAM:	PF00002;	7um_2;	1.		
DR	PRINTS:	PR00249;	GPCRSECRETIN.		
DR	PRINTS:	PR00393;	PTRHOMONER.		
DR	PROSITE:	PS00649;	G_PROTEIN_RECEP_F2_1; 1.		
DR	PROSITE:	PS00650;	G_PROTEIN_RECEP_F2_2; 1.		
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.				
FT	SIGNAL	1	POTENTIAL.		
FT	CHAIN	27	591		
FT	DMAIN	27	188		
FT	TRANSMEM	189	212		
FT	DMAIN	214	219		
FT	TRANSMEM	220	239		
FT	DMAIN	240	282		
FT	TRANSMEM	283	306		
FT	DMAIN	307	320		
FT	TRANSMEM	321	342		
FT	DMAIN	343	361		
FT	TRANSMEM	362	382		
FT	DMAIN	383	409		
FT	TRANSMEM	410	428		
FT	DMAIN	429	440		
FT	TRANSMEM	441	463		
FT	DMAIN	464	591		
FT	CARBOHYD	151	151		
FT	CARBOHYD	161	161		
FT	CARBOHYD	166	166		
FT	CARBOHYD	176	176		
SEQ	SEQUENCE	591 AA;	66260 MW; 21944F3051B9E9C1 CRC64;		
Query Match		48.3%;	Score 145; DB 1; Length 591;		
Best Local Similarity		57.9%;	Pred. No. 4.3e-10;		
Matches	22;	Conservative	9; Mismatches 7; Indels 0; Gaps 0;		
QY	1	DNHKGVAFRICNPNGTWDFMHSLNKTKWANYSDCLRFL	38		
		: : : : : : : : : :			
Db	137	DNHKGHAYRCDRNGSEVVPVGHNRWTWANYSECLAFM	174		
RESULT	7				
ID	PTHR_HUMAN	STANDARD;	PRT; 593 AA.		
AC	Q03431;				
DT	01-OCT-1993 (Rel. 27, Created)				
DT	01-OCT-1993 (Rel. 27, Last sequence update)				
DT	15-JUL-1999 (Rel. 38, Last annotation update)				
DE	PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR PRECURSOR (PTH/PTHr RECEPTOR).				
DE	PTHRI OR PTHR.				
GN	Homo sapiens (Human).				
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=KIDNEY;				
RX	MEDLINE: 93238641.				
RA	Schipani E., Karga H., Karaplis A.C., Potts J.T. Jr., Kronenberg H.M.,				
RA	About-Samra A.-B., Segre G.V., Jueppner H.;				
RT	"Identical complementary deoxyribonucleic acids encode a human renal				
RT	and bone parathyroid hormone (PTH)/PTH-related peptide receptor.";				
RL	Endocrinology 132:2157-2165(1993).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=KIDNEY;				
RX	MEDLINE: 93387403.				
RA	Schneider H., Feyen J.-H., Rao Movva N.;				
RT	"Cloning and functional expression of a human parathyroid hormone				
RT	receptor";				
RL	Eur. J. Pharmacol. 246:149-155(1993).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	MEDLINE: 95263723.				
RX					

RN RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/HEHA;
 RX MEDLINE; 95034305.
 RA Karperien M., van Dijk T.B., Hoelmakers T., Cremers F.,
 RA Abou-Samra A.-B., Boonstra J., de Laat S.W., Defize L.H.K.;
 RT "Expression pattern of parathyroid hormone/parathyroid hormone
 RT related peptide receptor mRNA in mouse postimplantation embryos
 RT indicates involvement in multiple developmental processes.";
 RL Mech. Dev. 47:29-42(1994).
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C;
 RX MEDLINE; 94255468.
 RA McCuaig K.A., Clarke J.C., White J.H.;
 RT "Molecular cloning of the gene encoding the mouse parathyroid
 RT hormone/parathyroid hormone-related peptide receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:5051-5055(1994).
 CC -1- FUNCTION: THIS IS A RECEPTOR FOR PARATHYROID HORMONE AND FOR
 CC PARATHYROID HORMONE-RELATED PEPTIDE. THE ACTIVITY OF THIS
 CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATES ADENYLYL
 CC CYCLASE AND ALSO A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER
 CC SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X78936; CAA55536.1; -
 CC EMBL; L34611; AAA40011.1; -
 CC EMBL; L34608; AAA40011.1; JOINED.
 CC EMBL; L34607; AAA40011.1; JOINED.
 CC EMBL; L34609; AAA40011.1; JOINED.
 CC EMBL; L34610; AAA40011.1; JOINED.
 CC GCRDB; GCR_1005; -
 CC GCRDB; GCR_1614; -
 CC MGD; MGI:97801; PTHR.
 CC INTERPRO; IPR000832; -
 CC INTERPRO; IPR002170; -
 CC PFAM; PF00002; 7tm.2; 1.
 CC PRINTS; PR00249; GPCRSECRETIN.
 CC PRINTS; PR00393; PTHORMONER.
 CC PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
 CC PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 591
 FT
 FT DOMAIN 27 188
 FT TRANSMEM 189 212
 FT DOMAIN 214 219
 FT TRANSMEM 220 239
 FT DOMAIN 240 282
 FT TRANSMEM 283 306
 FT DOMAIN 307 320
 FT TRANSMEM 321 342
 FT DOMAIN 343 361
 FT TRANSMEM 362 382
 FT DOMAIN 383 409
 FT TRANSMEM 410 428
 FT DOMAIN 429 440
 FT TRANSMEM 441 463
 FT DOMAIN 464 591
 FT CARBOHYD 151 151
 FT CARBOHYD 161 161
 FT CARBOHYD 166 166
 FT CARBOHYD 176 176

FT CONFLICT 27 28 DA -> TS (IN REF. 2).
 FT CONFLICT 464 465 MISSING (IN REF. 2).
 FT CONFLICT 500 501 GA -> VS (IN REF. 2).
 SQ SEQUENCE 591 AA; 66313 MW; F7876F8D388BDDDF CRC64;
 Query Match 48.3%; Score 145; DB 1; Length 591;
 Best Local Similarity 57.9%; Pred. No. 4.3e-10;
 Matches 22; Conservative 9; Mismatches 7; Indels 0; Gaps 0;
 QY 1 DFNHKGVAFRHCNPNGTWDFHSLNKTWNTYSDCLRFL 38
 Db 137 DFNHKGHAYRCDRNGSEVVPGHNRWTWNTYSECLKFM 174
 RESULT 6
 PTHR_RAT STANDARD; PRT; 591 AA.
 ID PTHR_RAT
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR
 DE PRECURSOR (PTH/PTHR RECEPTOR).
 GN PTHR1 OR PTHR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BONE;
 RX MEDLINE; 92212903.
 RA Abou-Samra A.-B., Jueppner H., Force T., Freeman M.W., Kong X.-F.,
 RA Schipani E., Urena P., Richards J., Bonventre J.V., Potts J.T. Jr.,
 RA Kronenberg H.M., Segre G.V.;
 RA "Expression cloning of a common receptor for parathyroid hormone and
 RT parathyroid hormone-related peptide from rat osteoblast-like cells: a
 RT single receptor stimulates intracellular accumulation of both cAMP
 RT and inositol trisphosphates and increases intracellular free
 RT calcium.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2732-2736(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 94292182.
 RA Pausova Z., Bourdon J., Clayton D., Mattei M.-C., Seldin M.F.,
 RA Janicic N., Riviere M., Szpirer J., Levan G., Szpirer C.;
 RA "Cloning of a parathyroid hormone/parathyroid hormone-related peptide
 RT receptor (PTHr) cDNA from a rat osteosarcoma (UMR 106) cell line:
 RT chromosomal assignment of the gene in the human, mouse, and rat
 RT genomes.";
 RL Genomics 20:20-26(1994).
 CC -1- FUNCTION: THIS IS A RECEPTOR FOR PARATHYROID HORMONE AND FOR
 CC PARATHYROID HORMONE-RELATED PEPTIDE. THE ACTIVITY OF THIS
 CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATES ADENYLYL
 CC CYCLASE AND ALSO A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER
 CC SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M77184; AAA41811.1; -
 CC EMBL; L19475; AAA68098.1; -
 CC GCRDB; GCR_0206; -
 CC GCRDB; GCR_0938; -
 CC INTERPRO; IPR000832; -
 CC INTERPRO; IPR002170; -

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 24
FT CHAIN 25
FT DOMAIN 26
FT TRANSMEM 146
FT TRANSMEM 170
FT TRANSMEM 177
FT TRANSMEM 197
FT TRANSMEM 238
FT TRANSMEM 261
FT TRANSMEM 276
FT TRANSMEM 298
FT TRANSMEM 317
FT TRANSMEM 337
FT TRANSMEM 363
FT TRANSMEM 384
FT TRANSMEM 395
FT TRANSMEM 418
FT TRANSMEM 501
FT CARBOHYD 106
FT CARBOHYD 116
FT CARBOHYD 121
SQ SEQUENCE 550 AA; 62235 MW; 2ADD14DBA68A9BF8 CRC64;

Query Match 96.3%; Score 289; DB 1; Length 550;
Best Local Similarity 98.0%; Pred. No. 2.4e-27;
Matches 50; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DFNHKGVAFRHCNPGTWDFHSLNKTWANYSDCLRFLOPDISIGKQEFCE 51
|||||
DB 92 DFNHKGVAFRHCNPGTWDFHSLNKTWANYSDCLRFLOPDISIGKQEFCE 142

RESULT 2
PTH2_RAT
ID PTH2_RAT STANDARD; PRT; 546 AA.
AC P70555;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PARATHYROID HORMONE RECEPTOR PRECURSOR (PTH2 RECEPTOR).
GN PTHR2
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RX MEDLINE; 96426194.
RA Usdin T.B., Bonner T.I., Harta G., Mezey E.;
RT "Distribution of parathyroid hormone-2 receptor messenger ribonucleic acid in rat."
RL Endocrinology 137:4285-4297(1996).
CC -!- FUNCTION: THIS IS A SPECIFIC RECEPTOR FOR PARATHYROID HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE CYCLASE. PTH2 MAY BE RESPONSIBLE FOR PTH EFFECTS IN A NUMBER OF PHYSIOLOGICAL SYSTEMS. IT MAY PLAY A SIGNIFICANT ROLE IN PANCREATIC FUNCTION. PTH2 PRESENCE IN NEURONS INDICATES THAT IT MAY FUNCTION AS A NEUROTRANSMITTER RECEPTOR.
CC -!- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN BRAIN, ARTERIAL AND CARDIAC ENDOTHELIUM. FOUND AS WELL IN SPERM, IN THE HEAD OF THE EPIDIDYMIS. LOWER EXPRESSION IS FOUND IN VASCULAR SMOOTH MUSCLE, EXOCRINE PANCREAS, TESTIS AND PLACENTA.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
CC

EMBL; U55836; AAC52849.1; --
DR GCRDB; GCR_1413; --
DR INTERPRO; IPR000832; --
DR PFAM; PF00002; 7tm.2; 1.
DR PRINTS; PR00249; GPCRSECRETIN
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 24
FT CHAIN 25
FT DOMAIN 27
FT TRANSMEM 144
FT TRANSMEM 167
FT TRANSMEM 174
FT TRANSMEM 194
FT TRANSMEM 235
FT TRANSMEM 236
FT TRANSMEM 259
FT TRANSMEM 274
FT TRANSMEM 296
FT TRANSMEM 314
FT TRANSMEM 334
FT TRANSMEM 335
FT TRANSMEM 362
FT TRANSMEM 381
FT TRANSMEM 392
FT TRANSMEM 414
FT TRANSMEM 415
FT CARBOHYD 106
FT CARBOHYD 116
FT CARBOHYD 121
SQ SEQUENCE 546 AA; 61800 MW; 2825AE4040313527 CRC64;

Query Match 82.0%; Score 246; DB 1; Length 546;
Best Local Similarity 84.3%; Pred. No. 3.3e-22;
Matches 43; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

QY 1 DFNHKGVAFRHCNPGTWDFHSLNKTWANYSDCLRFLOPDISIGKQEFCE 51
|||||
DB 92 DFNHKGVAFRHCNPGTWDFHSLNKTWANYSDCLRFLOPDISIGKQEFCE 140

RESULT 3
PTHR_DIDMA
ID PTHR_DIDMA STANDARD; PRT; 585 AA.
AC P25107;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR PRECURSOR (PTH/PTHR RECEPTOR).
GN PTHR.
OS Didelphis marsupialis virginiana (North American opossum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.
RN [1]
RX MEDLINE; 92054592.
RA Jueppner H., Abou-Samra A.-B., Freeman M., Kong X.-F., Schipani E., Richards J., Kolakowski L.F. Jr., Hock J., Potts J.T. Jr., Kronenberg H.M., Segre G.V.;
RA "A G protein-linked receptor for parathyroid hormone and parathyroid hormone-related peptide."
RL Science 254:1024-1026(1991).
CC -!- FUNCTION: THIS IS A RECEPTOR FOR PARATHYROID HORMONE AND FOR RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE CYCLASE AND ALSO A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC
CC PARATHYROID HORMONE-RELATED PEPTIDE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE CYCLASE AND ALSO A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 09:03:44 ; Search time 58.45 Seconds
(without alignments)
27.876 Million cell updates/sec

Title: US-09-236-468A-2_COPY_92_142

Perfect score: 300

Sequence: 1 DFNHKGVAFRHCPNGTWDF.....SDCLRFQLQPDISIGKQEFCE 51

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 87993

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	289	96.3	550	1 PTR2_HUMAN	P49190 homo sapien
2	246	82.0	546	1 PTH2_RAT	P70555 rattus norv
3	146	48.7	585	1 PTH2_HUMAN	P25107 didelphis m
4	145	48.3	585	1 PTH2_PIG	P50133 sus scrofa
5	145	48.3	591	1 PTH2_MOUSE	P41593 mus musculu
6	145	48.3	591	1 PTH2_RAT	P25961 rattus norv
7	145	48.3	593	1 PTH2_HUMAN	Q03431 homo sapien
8	76.5	25.5	516	1 CALR_RAT	P32214 rattus norv
9	75	25.0	515	1 CALR_MOUSE	Q60755 mus musculu
10	71	23.7	490	1 CALR_HUMAN	P30988 homo sapien
11	70	23.3	478	1 CALR_CAVPO	O08893 cavia porce
12	69	23.0	461	1 CRR2_HUMAN	Q16602 homo sapien
13	69	23.0	474	1 CALR_RABIT	P79222 oryctolagus
14	68.5	22.8	760	1 CO2_MOUSE	P21180 mus musculu
15	67.5	22.5	413	1 CRF2_XENLA	O42603 xenopus lae
16	64	21.3	464	1 CRR2_RAT	Q63118 rattus norv
17	60.5	20.2	498	1 CALR_PIG	P25117 sus scrofa
18	59.5	19.8	2504	1 FAS_HUMAN	P49327 homo sapien
19	59.5	19.8	2652	1 RRPB_IBVB	P26314 avian infec
20	59	19.7	499	1 ACH7_BOVIN	P54131 bos taurus
21	57.5	19.2	681	1 VGP_WABVP	P35253 marburg vir
22	57	19.0	504	1 SVK1_ECOLI	P13030 escherichia
23	56.5	18.8	530	1 PTH2_HUMAN	P06861 homo sapien
24	56.5	18.8	532	1 PTH2_HUMAN	P10696 homo sapien
25	56.5	18.8	535	1 PTH2_HUMAN	P05187 homo sapien
26	56.5	18.8	535	1 PTH2_HUMAN	P05188 homo sapien
27	56.5	18.8	681	1 VGP_WABVP	P35254 marburg vir
28	56.5	18.8	938	1 NMZ1_HUMAN	Q05586 homo sapien
29	56.5	18.8	938	1 NMZ1_MOUSE	P35438 mus musculu
30	56.5	18.8	938	1 NMZ1_RAT	P35439 rattus norv
31	55.5	18.5	752	1 CO2_HUMAN	P06681 homo sapien
32	55.5	18.5	913	1 PTH2_HUMAN	Q9UJU3 homo sapien
33	55	18.3	314	1 PPY_DROME	P11612 drosophila

RESULT 1

PTH2_HUMAN

ID PTR2_HUMAN STANDARD; PRT; 550 AA.

AC P49190;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE PARATHYROID HORMONE RECEPTOR PRECURSOR (PTH2 RECEPTOR).

GN PTHR2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]

RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.

RC TISSUE=BRAIN;

RX MEDLINE; 95318121.

RA Usdin T.B., Gruber C., Bonner T.I.;

RT "Identification and functional expression of a receptor selectively

RT recognizing parathyroid hormone, the PTH2 receptor.";

RL J. Biol. Chem. 270:15455-15458(1995).

RN [2]

RP SEQUENCE OF 26-40 AND 306-550 FROM N.A.

RX MEDLINE; 97079671.

RA Usdin T.B., Modi W., Bonner T.I.;

RT "Assignment of the human PTH2 receptor gene (PTH2) to chromosome 2q33

RT by fluorescence in situ hybridization.";

RL Genomics 37:140-141(1996)

CC -!- FUNCTION: THIS IS A SPECIFIC RECEPTOR FOR PARATHYROID HORMONE. THE

CC ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE

CC ADENYL CYCLASE.

CC -!- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN BRAIN AND PANCREAS.

CC ALSO EXPRESSED IN THE TESTIS.

CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U25128; AAC50157.1; -

DR EMBL; U47124; AAA96796.1; -

DR EMBL; U47129; AAC50767.1; -

DR EMBL; U47125; AAC50767.1; JOINED.

DR EMBL; U47126; AAC50767.1; JOINED.

DR EMBL; U47127; AAC50767.1; JOINED.

DR EMBL; U47128; AAC50767.1; JOINED.

DR GCRDB; GCR.2003; -

DR MIM; 601469; -

DR INTERPRO; IPR000832; -

DR PFAM; PF00002; 7tm2; 1.

DR PRINTS; PR00249; GPCRSECRETIN.

DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.

DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.

ALIGNMENTS

34	55	18.3	324	1	CATV_NPVCF	P41715 choristoneu
35	55	18.3	502	1	ACH7_MOUSE	P49582 mus musculu
36	55	18.3	502	1	ACH7_RAT	Q05941 rattus norv
37	55	18.3	1159	1	DP3A_HAEIN	P43743 haemophilus
38	54.5	18.2	411	1	CRF2_HUMAN	Q13324 homo sapien
39	54.5	18.2	431	1	CRF2_MOUSE	Q60748 mus musculu
40	54.5	18.2	528	1	PPB1_HUMAN	P09923 homo sapien
41	54	18.0	512	1	RCK1_YEAST	P39622 saccharomyc
42	54	18.0	1036	1	P200_MYCPN	P75211 mycoplasma
43	53.5	17.8	411	1	CRF2_RAT	P47866 rattus norv
44	53.5	17.8	415	1	CRFR_MOUSE	P35347 mus musculu
45	53.5	17.8	415	1	CRFR_RAT	P35353 rattus norv

THIS PAGE BLANK (USPTO)

I60800
calcitonin receptor clb precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 11-Jan-2000
C:Accession: I60800; S33747
R:Sexton, P.M.; Housami, S.; Hilton, J.M.; O'Keefe, L.M.; Center, R.J.; Gillespie,
Mol. Endocrinol. 7, 815-821, 1993
A:Title: Identification of brain isoforms of the rat calcitonin receptor.
A:Reference number: A37430; MUID:93368608
A:Accession: I60800
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-515 <RES>
A:Cross-references: GB:I13040; NID:Q294532; PIDN:AAA03031.1; PID:Q294533
R:Albrandt, K.; Mull, E.; Brady, E.M.G.; Herich, J.; Moore, C.X.; Beaumont, K.
FEBS Lett. 325, 225-232, 1993
A:Title: Molecular cloning of two receptors from rat brain with high affinity for sal
A:Reference number: S33746; MUID:93307500
A:Accession: S33747
A:Molecule type: mRNA
A:Residues: 1-147,'L',149-458,'R',459-477,'L',479-515 <ALB>
A:Cross-references: GB:I14618; NID:G347431; PIDN:AAA65965.1; PID:G347432
C:Superfamily: glucagon receptor
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 25.5%; Score 76.5; DB 2; Length 515;
Best Local Similarity 40.5%; Pred. No. 0.11;
Matches 17; Conservative 7; Mismatches 15; Indels 3; Gaps 3;

QY 1 DFNHKGVAFRHCNPNGTWDPMH-SLKNTWANYSDCLRFLOPD 41
||| :|| :|| :|| :|| :|| :|| :||
Db 101 DFDPTEKYSKYCDENGSEW-FRHFDPSNRFTWSNYTLCSNAF-TPD 140

```

RESULT 10
I49154
calcitonin receptor lb - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Nov-1999
C:Accession: I49154
R:Yamin, M.; Gorn, A.H.; Flannery, M.R.; Jenkins, N.A.; Gilbert, D.J.; Copeland, N.G.
Endocrinology 135, 2635-2643, 1994
A:Title: Cloning and characterization of a mouse brain calcitonin receptor complement
A:Reference number: I49154; MUID:95080136
A:Accession: I49154
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-515 <RES>
A:Cross-references: EMBL:U18542; NID:g604510; PIDN:AAA69521.1; PID:g604511
C:Superfamily: glucagon receptor

Query Match          25.0%; Score 75; DB 2; Length 515;
Best Local Similarity 39.5%; Pred. No. 0.16;
Matches 15; Conservative 7; Mismatches 14; Indels 2; Gaps 2;

Qy 1 DFNHKGVAFRHCNPNGTWDPMH-SLNKWTWANYSDCLRF 37
   ||| :|| | | | | | | | | | | | | | | | |
Db 101 DFDTAEKVSKYCDENGEW-FRRHPDSNRRTWSNYTLCNAF 137

```

```

RESULT 11
I37217
calcitonin receptor - human
C:Species: Homo sapiens (man)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 11-Jan-2000
C:Accession: I37217; S43673; S44209
R:Kuestner, R.E.; Eliod, R.D.; Grant, F.J.; Hagen, F.S.; Kuijper, J.L.; Matthews, S.
n, P.M.; Moore, E.E.
Mol. Pharmacol.. 46, 246-255, 1994
A:Title: Cloning and characterization of an abundant subtype of the human calcitonin
A:Reference number: I37217; MUID:94359487

```

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model
Run on: November 8, 2000, 08:53:22 ; Search time 99,87 Seconds
(without alignments)
32.407 Million cell updates/sec

Title: US-09-236-468a-2_COPY_92_142
Perfect score: 300
Sequence: 1 DFNHKGVAFRHCNPNGTWF.....SDCLRFLOPDISIGKQEFCE 51

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 182106

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.65:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	289	96.3	550	2 A57519	parathyroid hormon-
2	146	48.7	585	2 A39286	parathyroid hormon
3	145	48.3	589	2 I59297	parathyroid hormon
4	145	48.3	591	2 S44203	parathyroid hormon
5	145	48.3	591	2 I54195	parathyroid hormon
6	145	48.3	593	2 A49191	parathyroid hormon
7	76.5	25.5	478	2 A37430	calcitonin recepto
8	76.5	25.5	479	2 S33746	calcitonin recepto
9	76.5	25.5	515	2 I60800	calcitonin recepto
10	75	25.0	515	2 I49154	calcitonin recepto
11	71	23.7	474	2 I37217	calcitonin recepto
12	71	23.7	490	2 S34486	calcitonin recepto
13	69	23.0	461	2 JC2477	calcitonin recepto
14	68.5	22.8	760	1 C2MS	classical-compleme
15	67	22.3	591	2 S43428	omega-crystallin -
16	64	21.3	462	2 I60194	calcitonin-like re
17	60.5	20.2	484	2 A39285	calcitonin recepto
18	60.5	20.2	498	2 I47130	calcitonin recepto
19	59.5	19.8	2504	1 A57788	fatty-acid synthas
20	59.5	19.8	2652	1 VFTHB2	genome polypeptid
21	59	19.7	274	2 E81319	probable glucosylt
22	57.5	19.2	681	2 A45705	type I transmembra
23	57.5	19.2	811	2 S39901	nwsa protein - Bra
24	57	19.0	505	1 SVCKT	lysine--trna ligas
25	56.5	18.8	221	2 I51308	N-methyl-D-asparta
26	56.5	18.8	532	2 S12076	alkaline phosphata
27	56.5	18.8	535	1 PAHUA	alkaline phosphata
28	56.5	18.8	614	2 S43427	intermediate filam
29	56.5	18.8	681	2 S33316	structural protein

30 56.5 18.8 885 2 JN0339 N-methyl-D-asparta
31 56.5 18.8 901 2 JN0337 N-methyl-D-asparta
32 56.5 18.8 906 2 A46296 N-methyl-D-asparta
33 56.5 18.8 906 2 JN0341 N-methyl-D-asparta
34 56.5 18.8 922 2 JN0340 N-methyl-D-asparta
35 56.5 18.8 922 2 JN0338 N-methyl-D-asparta
36 56.5 18.8 938 2 S19710 N-methyl-D-asparta
37 56.5 18.8 938 2 A46612 N-methyl-D-asparta
38 56.5 18.8 938 2 S21104 N-methyl-D-asparta
39 56.5 18.8 943 2 A47551 N-methyl-D-asparta
40 56.5 18.8 959 2 JN0336 N-methyl-D-asparta
41 56.5 18.8 965 2 I51244 pol protein - bovi
42 56 18.7 852 2 S29358 complement C2 prec
43 55.5 18.5 752 1 C2HU phosphoprotein pho
44 55 18.3 314 1 PAFFV cathepsin - Choris
45 55 18.3 324 2 S62735

ALIGNMENTS

RESULT 1

A57519
parathyroid hormone receptor 2 precursor - human
N:Alternate names: PTH2 receptor
C:Species: Homo sapiens (man)
C:Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 24-Nov-1999
C:Accession: A57519
R:Usdin, T.B.; Gruber, C.; Bonner, T.I.
J. Biol. Chem. 270, 15455-15458, 1995
A:Title: Identification and functional expression of a receptor selectively recognizi
A:Reference number: A57519; MUID:95318121
A:Accession: A57519
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-550 <USD>
A:Cross-references: GB:U25128; MID:g887966; PIDN:AAC50157.1; PID:g887967
C:Genetics:
A:Gene: GDB:PTH2; PTHR2
A:Cross-references: GDB:731977; OMIM:601469
A:Map position: 2q33-2q33
C:Superfamily: glucagon receptor
C:Keywords: hormone receptor

Query Match 96.3%; Score 289; DB 2; Length 550;
Best Local Similarity 98.0%; Pred. No. 2e-27; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 1;

Oy 1 DFNHKGVAFRHCNPNGTWFHSLNKTWANSCLRFLOPDISIGKQEFCE 51
|||||
Db 92 DFNHKGVAFRHCNPNGTWFHSLNKTWANSCLRFLOPDISIGKQEFCE 142

RESULT 2

A39286
parathyroid hormone / parathyroid hormone-related peptide - North American opossum
C:Species: Didelphis virginiana, Didelphis marsupialis virginiana (North American opo
C:Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 11-Jan-2000
R:Jueppner, H.; Abou-Samra, A.B.; Freeman, M.; Kong, X.F.; Schipani, E.; Richards, J.
Science 254, 1024-1026, 1991
C:Accession: A39286
A:Title: A G protein-linked receptor for parathyroid hormone and parathyroid hormone-
A:Reference number: A39286; MUID:92054592
A:Accession: A39286
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-585 <JUE>
A:Cross-references: GB:M74445
C:Superfamily: glucagon receptor
C:Keywords: G protein-coupled receptor; transmembrane protein

THIS PAGE BLANK (USPTO)

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/681,702
 ; FILING DATE: 04-MAY-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fraser, Janis K.
 ; REGISTRATION NUMBER: 34,819
 ; REFERENCE/DOCKET NUMBER: 00786/071003
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 591 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-468-249A-20

Query Match 48.3%; Score 145; DB 2; Length 591;
 Best Local Similarity 57.9%; Pred. No. 9,7e-12;
 Matches 22; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 1 DFNHKGVAFRHCNPNGTWDFMHSLNKNTWANYSDCLRL 38
 Db 137 DFNHKGHAYRRCDRNGSWELVPGHNRTWANYSECLKFM 174

RESULT 10
 US-08-468-249A-21
 ; Sequence 21, Application US/08468249A
 ; Patent No. 5886148
 ; GENERAL INFORMATION:
 ; APPLICANT: Segre et al., Gino V.
 ; TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR AND DNA
 ; TITLE OF INVENTION: ENCODING SAME
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/468,249A
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/864,475
 ; FILING DATE: 06-APR-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/681,702
 ; FILING DATE: 04-MAY-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fraser, Janis K.
 ; REGISTRATION NUMBER: 34,819
 ; REFERENCE/DOCKET NUMBER: 00786/071003
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906
 ; INFORMATION FOR SEQ ID NO: 21:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 593 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-468-249A-21

Query Match 48.3%; Score 145; DB 2; Length 593;
 Best Local Similarity 57.9%; Pred. No. 9,8e-12;
 Matches 22; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 1 DFNHKGVAFRHCNPNGTWDFMHSLNKNTWANYSDCLRL 38
 Db 137 DFNHKGHAYRRCDRNGSWELVPGHNRTWANYSECVKFL 174

RESULT 11
 US-08-468-011A-16
 ; Sequence 16, Application US/08468011A
 ; Patent No. 6030804
 ; GENERAL INFORMATION:
 ; APPLICANT: Soppet, Daniel R
 ; APPLICANT: Yi, Li
 ; APPLICANT: Rosen, Craig A
 ; APPLICANT: Ruben, Steven
 ; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
 ; TITLE OF INVENTION: H1TDG74
 ; NUMBER OF SEQUENCES: 28
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
 ; ADDRESSEE: Stewart & Olstein
 ; STREET: 6 Becker Farm Road
 ; CITY: Roseland
 ; STATE: NJ
 ; COUNTRY: USA
 ; ZIP: 07068-1739
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 INCH DISKETTE
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WORD PERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/468,011A
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MULLINS, J.G.
 ; REGISTRATION NUMBER: 33,073
 ; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-994-1700
 ; TELEFAX: 201-994-1744
 ; INFORMATION FOR SEQ ID NO: 16:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 60 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-468-011A-16

Query Match 33.7%; Score 101; DB 3; Length 60;
 Best Local Similarity 53.6%; Pred. No. 5,9e-07;
 Matches 15; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 DFNHKGVAFRHCNPNGTWDFMHSLNKNTW 28
 Db 33 DFNHKGHAYRRCDRNGSWELVPGNNRTW 60

RESULT 12
 PCT-US94-09235-2
 ; Sequence 2, Application PC/TUS9409235
 ; GENERAL INFORMATION:
 ; APPLICANT: LI, ET AL.
 ; TITLE OF INVENTION: Calcitonin Receptor
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:

TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 515 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-468-249A-18

Query Match 48.7%; Score 146; DB 2; Length 515;
Best Local Similarity 57.9%; Pred. No. 6.le-12;
Matches 22; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 1 DENHKGVAFRHCNPCTWDFMHSLNKNTWANYSDCLRL 38
DB 134 DENHKGVAFRHCNPCTWDFMHSLNKNTWANYSDCLRL 171

RESULT 5

US-08-142-439A-6
Sequence 6, Application US/08142439A
Patent No. 5670360
GENERAL INFORMATION:
APPLICANT: Thorens, Bernard
TITLE OF INVENTION: Receptor for the Glucagon-Like-Peptide-1
TITLE OF INVENTION: (GLP-1)
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5670360 No. 5670360disk of No. 5670360th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142.439A
FILING DATE: 24-NOV-93
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DK 398/92
FILING DATE: 25-MAR-92
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP93/00697
FILING DATE: 23-MAR-93
ATTORNEY/AGENT INFORMATION:
NAME: Harrington, James J.
REGISTRATION NUMBER: 38,711
REFERENCE/DOCKET NUMBER: 3756.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Didelphis virginiana
US-08-142-439A-6

Query Match 48.7%; Score 146; DB 1; Length 585;

Best Local Similarity 57.9%; Pred. No. 7e-12;
Matches 22; Conservative 9; Mismatches 7; Indels 0; Gaps 0;
QY 1 DENHKGVAFRHCNPCTWDFMHSLNKNTWANYSDCLRL 38
DB 134 DENHKGVAFRHCNPCTWDFMHSLNKNTWANYSDCLRL 171

RESULT 6

US-08-142-551B-125
Sequence 125, Application US/08142551B
Patent No. 5814603
GENERAL INFORMATION:
APPLICANT: Oldenburg, Kevin R.
APPLICANT: Selick, Harold E.
TITLE OF INVENTION: COMPOUNDS WITH PTH ACTIVITY AND
TITLE OF INVENTION: RECOMBINANT DNA VECTORS ENCODING SAME
NUMBER OF SEQUENCES: 132
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: 699 Prince Street
CITY: Alexandria
STATE: Virginia
COUNTRY: US
ZIP: 22313

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142.551B
FILING DATE: 25-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/077,296
FILING DATE: 14-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/898,219
FILING DATE: 12-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,677
FILING DATE: 22-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Swiss, Gerald F.
REGISTRATION NUMBER: 30,113
REFERENCE/DOCKET NUMBER: 000324-010
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 854-7400
TELEFAX: (415) 854-8275
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..585
OTHER INFORMATION: /note= "PTH receptor"
US-08-142-551B-125

Query Match 48.7%; Score 146; DB 2; Length 585;
Best Local Similarity 57.9%; Pred. No. 7e-12;
Matches 22; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 1 DENHKGVAFRHCNPCTWDFMHSLNKNTWANYSDCLRL 38
DB 134 DENHKGVAFRHCNPCTWDFMHSLNKNTWANYSDCLRL 171

RESULT 7

Db 92 DFNHKGVAFRHCNPNGTWDFMHSLNKWTWANSYSDCLRFLOPDISIGKQEFCE 142

RESULT 2
PCT-US95-07085-2
; Sequence 2, Application PC/TUS9507085
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone Receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07085
; FILING DATE: 05-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-393
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-07085-2

Query Match 100.0%; Score 300; DB 4; Length 541;
Best Local Similarity 100.0%; Pred. No. 9.3e-33; Mismatches 0; Indels 0; Gaps 0;
Matches 51; Conservative 0;

Qy 1 DFNHKGVAFRHCNPNGTWDFMHSLNKWTWANSYSDCLRFLOPDISIGKQEFCE 51
|||||
Db 92 DFNHKGVAFRHCNPNGTWDFMHSLNKWTWANSYSDCLRFLOPDISIGKQEFCE 142

RESULT 3
US-08-468-011A-15
; Sequence 15, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland

STATE: NJ
COUNTRY: USA
ZIP: 07068-1739
COMPUTER-READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,011A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-468-011A-15

Query Match 58.0%; Score 174; DB 3; Length 60;
Best Local Similarity 100.0%; Pred. No. 7.8e-17; Mismatches 0; Indels 0; Gaps 0;
Matches 28; Conservative 0;

Qy 1 DFNHKGVAFRHCNPNGTWDFMHSLNKWTW 28
|||||
Db 33 DFNHKGVAFRHCNPNGTWDFMHSLNKWTW 60

RESULT 4
US-08-468-249A-18
; Sequence 18, Application US/08468249A
; Patent No. 5886148
; GENERAL INFORMATION:
; APPLICANT: Segre et al., Gino V.
; TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR AND DNA
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,249A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,475
; FILING DATE: 06-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,702
; FILING DATE: 04-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/071003
; TELECOMMUNICATION INFORMATION:

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:49:16 ; Search time 97.15 seconds
(without alignments)
8.799 Million cell updates/sec

Title: US-09-236-468A-2_COPY_92_142
Perfect score: 300
Sequence: 1 DFNHKGVAFRHCPNPTWDF.....SDCLRFLOPDISIGKQFCE 51

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 164575

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	300	100.0	541	3	US-08-468-011A-2
2	300	100.0	541	4	PCT-US95-07085-2
3	174	58.0	60	3	US-08-468-011A-15
4	146	48.7	515	2	US-08-468-249A-18
5	146	48.7	585	1	US-08-142-439A-6
6	146	48.7	585	2	US-08-142-551B-125
7	146	48.7	585	2	US-08-869-477-6
8	146	48.7	585	2	US-08-468-249A-19
9	145	48.3	591	2	US-08-468-249A-20
10	145	48.3	593	2	US-08-468-249A-21
11	101	33.7	60	3	US-08-468-011A-16
12	72	24.0	472	4	PCT-US94-09235-2
13	71	23.7	180	1	US-08-453-742-25
14	71	23.7	180	1	US-08-454-464-25
15	71	23.7	180	1	US-08-453-222-25
16	71	23.7	180	1	US-08-452-802-25
17	71	23.7	474	1	US-08-453-742-2
18	71	23.7	474	1	US-08-454-464-2
19	71	23.7	474	1	US-08-453-222-2
20	71	23.7	474	1	US-08-452-802-2
21	71	23.7	477	1	US-08-453-742-27
22	71	23.7	477	1	US-08-454-464-27
23	71	23.7	477	1	US-08-453-222-27
24	71	23.7	477	1	US-08-452-802-27
25	70	23.3	553	3	US-08-845-546-12
26	69	23.0	462	3	US-09-238-796-2
27	69	23.0	509	3	US-08-845-546-2
28	60.5	20.2	482	1	US-07-792-885A-1

29	60.5	20.2	482	1	US-08-142-439A-7	Sequence 7, Appli
30	60.5	20.2	482	2	US-08-869-477-7	Sequence 7, Appli
31	56.5	18.8	489	2	US-08-752-307B-3	Sequence 3, Appli
32	56.5	18.8	506	3	US-08-867-352-21	Sequence 21, Appli
33	56.5	18.8	530	2	US-08-752-307B-2	Sequence 2, Appli
34	56.5	18.8	777	2	US-08-231-193A-16	Sequence 16, Appli
35	56.5	18.8	777	2	US-08-486-273A-16	Sequence 16, Appli
36	56.5	18.8	777	2	US-08-480-474-16	Sequence 16, Appli
37	56.5	18.8	777	3	US-08-940-086A-16	Sequence 16, Appli
38	56.5	18.8	854	2	US-08-231-193A-32	Sequence 32, Appli
39	56.5	18.8	854	2	US-08-486-273A-32	Sequence 32, Appli
40	56.5	18.8	854	3	US-08-480-474-32	Sequence 32, Appli
41	56.5	18.8	854	3	US-08-940-086A-32	Sequence 32, Appli
42	56.5	18.8	870	2	US-08-231-193A-30	Sequence 30, Appli
43	56.5	18.8	870	2	US-08-486-273A-30	Sequence 30, Appli
44	56.5	18.8	870	3	US-08-480-474-30	Sequence 30, Appli
45	56.5	18.8	870	3	US-08-940-086A-30	Sequence 30, Appli

ALIGNMENTS

RESULT 1
US-08-468-011A-2
; Sequence 2, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-011A-2

Query Match 100.0%; Score 300; DB 3; Length 541;
Best Local Similarity 100.0%; Pred. No. 9,3e-33;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DFNHKGVAFRHCPNPTWDFMHSLNKTWYSDCLRFLOPDISIGKQFCE 51
|||||

SQ Sequence 477 AA;

Query Match 25.5%; Score 76.5; DB 15; Length 477;
Best Local Similarity 40.5%; Pred. No. 0.065;
Matches 17; Conservative 7; Mismatches 15; Indels 3; Gaps 3;

Qy 1 DFNHKGVAFRHCNPNGTWDEMH-SLNKWTWANSDCLEFLQPD 41
Db 99 dfdptekvskydengew-frhpsnrtsnyticnaf-tpd 138

RESULT 15

R62447 ID R62447 standard; protein; 515 AA.

AC R62447;

DT 21-JUN-1995 (first entry)

DE Rat calcitonin receptor Cib protein.

KW Calcitonin receptor Cib.

OS Rattus rattus.

FH Key Location/Qualifiers

FT Domain 1..144 /note= "extracellular portion"
FT Domain 212..256 /note= "extracellular portion"
FT Domain 315..333 /note= "extracellular portion"
FT Domain 400..410 /note= "extracellular portion"
FT Domain 145..168 /note= "extracellular portion"
FT Domain 185..211 /note= "transmembrane portion"
FT Domain 257..280 /note= "transmembrane portion"
FT Domain 293..314 /note= "transmembrane portion"
FT Domain 334..356 /note= "transmembrane portion"
FT Domain 380..399 /note= "transmembrane portion"
FT Domain 411..430 /note= "transmembrane portion"
FT Domain 169..184 /note= "intracellular portion"
FT Domain 281..292 /note= "intracellular portion"
FT Domain 357..379 /note= "intracellular portion"
FT Domain 431..515 /note= "intracellular portion"

PN WO9421665-A.

PD 29-SEP-1994.

PF 24-MAR-1994; 94WO-US03248.

PR 24-MAR-1993; 93US-0036784.

PA (AMYL-) AMYLIN PHARM INC.

PI Albrandt K, Beaumont K;

XX WPI; 1994-316927/39.

XX New calcitonin receptor proteins - used to identify agents for

PT treating e.g. obesity, anorexia, pain, diabetes mellitus or
PT insulin resistance
XX
PS Disclosure; Fig. 1; 59pp; English.
XX
CC The receptor protein can be used for determining the presence or
CC amount of or separating Cib receptor binding compounds in a
CC sample. They can also be used for producing antibodies. The
CC receptor protein is especially used for identifying calcitonin.
CC amylin or CGRP agonists or antagonists for treating conditions such
CC as obesity, anorexia, pain, diabetes mellitus impaired glucose
CC tolerance or insulin resistance.
XX
SQ Sequence 515 AA;

Query Match 25.5%; Score 76.5; DB 15; Length 515;
Best Local Similarity 40.5%; Pred. No. 0.071;
Matches 17; Conservative 7; Mismatches 15; Indels 3; Gaps 3;

Qy 1 DFNHKGVAFRHCNPNGTWDEMH-SLNKWTWANSDCLEFLQPD 41
Db 99 dfdptekvskydengew-frhpsnrtsnyticnaf-tpd 138

Search completed: November 8, 2000, 08:51:41
Job time: 253 sec

CC expressing the receptor can be used for diagnostic measurement of
 CC PTH serum levels.

SQ Sequence 593 AA;

Query Match 48.3%; Score 145; DB 17; Length 593;
 Best Local Similarity 57.9%; Pred. No. 1.6e-10;
 Matches 22; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

Oy 1 DFNHKGVAFRHCNPNGTWDEHSLNKTWANYSDCLREL 38
 Db 137 dlnhkgayrcdrngswelpvghnrtwanysecvklf 174

RESULT 13

W73317
 ID W73317 standard; Protein; 593 AA.

XX AC W73317;

XX DT 08-FEB-1999 (first entry)

XX DE Human Parathyroid hormone receptor.

XX KW Parathyroid hormone receptor; PTH receptor; antibody; therapy;
 KW PTH-related hypercalcaemia; human.

XX OS Homo sapiens.

XX PN US5840853-A.

XX PD 24-NOV-1998.

XX PF 06-JUN-1995; 95US-0471494.

XX PR 06-APR-1992; 92US-0864475.

XX PR 05-APR-1991; 91US-0681702.

XX PR 06-JUN-1995; 95US-0471494.

XX PA (GENO) GEN HOSPITAL CORP.

XX PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;

XX PI Schipani E, Segre GV;

XX DR WPI: 1999-034124/03.

XX DR N-PSDB; V08391.

XX PT Antibody to parathyroid hormone receptor - for diagnostic or
 PT therapeutic use

XX PS Claim 7; Fig 6; 63pp; English.

XX CC This sequence represents the human parathyroid hormone (PTH) receptor
 CC which is targeted by the antibody of the invention. The antibody
 CC of the invention is immunoreactive with naturally occurring human, rat or
 CC opossum PTH receptor. The antibody is useful for treating disorders
 CC characterised by overstimulation of PTH receptors by their ligand and for
 CC the diagnosis of PTH-related hypercalcaemia.

XX SQ Sequence 593 AA;

Query Match 48.3%; Score 145; DB 20; Length 593;
 Best Local Similarity 57.9%; Pred. No. 1.6e-10;
 Matches 22; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

Oy 1 DFNHKGVAFRHCNPNGTWDEHSLNKTWANYSDCLREL 38
 Db 137 dlnhkgayrcdrngswelpvghnrtwanysecvklf 174

RESULT 14

R62448
 ID R62448 standard; protein; 477 AA.
 XX AC R62448;
 XX DT 21-JUN-1995 (first entry)
 XX DE Rat calcitonin receptor C1a protein.
 XX KW Calcitonin receptor C1a.
 XX OS Rattus rattus.
 XX FH Key Location/Qualifiers
 FT Domain 1..144
 FT FT /note= "extracellular portion"
 FT Domain 212..219
 FT FT /note= "extracellular portion"
 FT Domain 278..296
 FT FT /note= "extracellular portion"
 FT Domain 363..373
 FT FT /note= "extracellular portion"
 FT Domain 145..168
 FT FT /note= "transmembrane portion"
 FT Domain 185..211
 FT FT /note= "transmembrane portion"
 FT Domain 220..243
 FT FT /note= "transmembrane portion"
 FT Domain 256..277
 FT FT /note= "transmembrane portion"
 FT Domain 297..319
 FT FT /note= "transmembrane portion"
 FT Domain 343..362
 FT FT /note= "transmembrane portion"
 FT Domain 374..393
 FT FT /note= "transmembrane portion"
 FT Domain 169..184
 FT FT /note= "intracellular portion"
 FT Domain 244..255
 FT FT /note= "intracellular portion"
 FT Domain 320..342
 FT FT /note= "intracellular portion"
 FT Domain 394..477
 FT FT /note= "intracellular portion"
 XX PN W09421665-A.
 XX PD 29-SEP-1994.
 XX PF 24-MAR-1994; 94WO-US03248.
 XX PR 24-MAR-1993; 93US-0036784.
 XX PA (AMYL-) AMYLIN PHARM INC.
 XX PI Albrandt K, Beaumont K;
 XX WPI: 1994-316927/39.
 XX CC New calcitonin receptor proteins - used to identify agents for
 XX PT treating e.g. obesity, anorexia, pain, diabetes mellitus or
 XX PT insulin resistance
 XX PS Disclosure; Fig. 1; 59pp; English.
 XX CC The receptor protein can be used for determining the presence or
 XX CC amount of or separating C1b receptor binding compounds in a
 XX CC sample. They can also be used for producing antibodies. The
 XX CC receptor protein is especially used for identifying calcitonin,
 XX CC amylin or CGRP agonists or antagonists for treating conditions such
 XX CC as obesity, anorexia, pain, diabetes mellitus impaired glucose
 XX CC tolerance or insulin resistance.


```

XX PR 06-APR-1992; 92US-0864475.
XX PR 05-APR-1991; 91US-0681702.
XX PA (GEO ) GEN HOSPITAL CORP.
XX PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
XX PI Schipani E, Segre GV;
XX DR WPI; 1996-139028/14.
XX DR N-PSDB; T15947.
XX PT DNA encoding vertebrate parathyroid hormone receptor - useful for
XX PT diagnosis and treatment of e.g. hypercalcaemia, hypo-calcaemia,
XX PT cancer etc.
XX PS Claim 1; Fig 3A-3E; 64pp; English.
XX CC A rat parathyroid hormone/parathyroid hormone-related protein
XX CC (PTH/PTHrP) receptor (R92277) is encoded by cDNA clone R15B
XX CC (T15947) isolated from a rat osteosarcoma ROS 17/2.8 cDNA library.
XX CC The receptor is a G-protein linked receptor having 7 transmembrane
XX CC domains. It induces an increase in intracellular cAMP and calcium
XX CC upon challenge with PTH or PTHrP. Recombinant receptor can be
XX CC produced in vector/host cell systems and used in the treatment,
XX CC diagnosis or prognosis of e.g. hypercalcaemia and hypocalcaemia,
XX CC to screen for (ant)agonists and to raise antibodies. Host cells
XX CC expressing the receptor are used for diagnostic measurement of PTH
XX CC serum levels.
XX SQ Sequence 591 AA;

Query Match 48.3%; Score 145; DB 17; Length 591;
Best Local Similarity 57.9%; Pred. No. 1.6e-10;
Matches 22; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

Qy 1 DFNHKGVAFRHCNPNGTWDFMHSLNKWTWANYSDCLRFL 38
| | | | | : : : : : : : : : : : : : : : : : : :
Db 137 dfnhkghayrrcdngswewvpgghrtwanyseclkm 174

RESULT 11
W73316
ID W73316 standard; Protein; 591 AA.
XX AC W73316;
XX DT 08-FEB-1999 (first entry)
XX DE Parathyroid hormone receptor R15B.
XX KW Parathyroid hormone receptor; PTH receptor; antibody; therapy;
XX KW PTH-related hypercalcaemia; rat.
XX OS Rattus sp.
XX PN US5840853-A.
XX PD 24-NOV-1998.
XX PF 06-JUN-1995; 95US-0471494.
XX PR 06-APR-1992; 92US-0864475.
XX PR 05-APR-1991; 91US-0681702.
XX PR 06-JUN-1995; 95US-0471494.
XX PA (GEO ) GEN HOSPITAL CORP.
XX PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
XX PI Schipani E, Segre GV;
XX PR WPI; 1999-034124/03.

```

```

DR N-PSDB; V08390.
XX XX Antibody to parathyroid hormone receptor - for diagnostic or
XX PT therapeutic use
XX PS Claim 6; Fig 3; 63pp; English.
XX CC This sequence represents the rat parathyroid hormone (PTH) receptor
XX CC R15B, which is targeted by the antibody of the invention. The antibody
XX CC of the invention is immunoreactive with naturally occurring human, rat or
XX CC opossum PTH receptor. The antibody is useful for treating disorders
XX CC characterised by overstimulation of PTH receptors by their ligand and for
XX CC the diagnosis of PTH-related hypercalcaemia.
XX SQ Sequence 591 AA;

Query Match 48.3%; Score 145; DB 20; Length 591;
Best Local Similarity 57.9%; Pred. No. 1.6e-10;
Matches 22; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

Qy 1 DFNHKGVAFRHCNPNGTWDFMHSLNKWTWANYSDCLRFL 38
| | | | | : : : : : : : : : : : : : : : : : : :
Db 137 dfnhkghayrrcdngswewvpgghrtwanyseclkm 174

RESULT 12
R92278
ID R92278 standard; Protein; 593 AA.
XX AC R92278;
XX DT 18-MAY-1996 (first entry)
XX DE Human kidney PTH/PTHrP receptor.
XX KW Parathyroid hormone; receptor; parathormone; PTH;
XX KW parathyroid hormone-related protein; PTHrP; calcium; homeostasis;
XX KW hypercalcaemia; hypocalcaemia; cancer.
XX OS Homo sapiens.
XX PN US5494806-A.
XX PD 27-FEB-1996.
XX PF 05-APR-1991; 91US-0681702.
XX PR 06-APR-1992; 92US-0864475.
XX PR 05-APR-1991; 91US-0681702.
XX PA (GEO ) GEN HOSPITAL CORP.
XX PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
XX PI Schipani E, Segre GV;
XX DR WPI; 1996-139028/14.
XX DR N-PSDB; T15948.
XX PT DNA encoding vertebrate parathyroid hormone receptor - useful for
XX PT diagnosis and treatment of e.g. hypercalcaemia, hypo-calcaemia,
XX PT cancer etc.
XX PS Claim 1; Fig 6A-6G; 64pp; English.
XX CC A human parathyroid hormone/parathyroid hormone-related protein
XX CC (PTH/PTHrP) receptor (R92278) is encoded by cDNA clone HK-1
XX CC (T15948) isolated from a human kidney cDNA library. The receptor
XX CC induces an increase in intracellular cAMP and intracellular free
XX CC calcium when challenged by PTH or PTHrP. Recombinant receptor can
XX CC be produced in vector/host cell systems and used in the treatment,
XX CC diagnosis or prognosis of e.g. hypercalcaemia and hypocalcaemia, to
XX CC screen for (ant)agonists and to raise antibodies. Host cells

```

In the blood. Cpds. capable of competing with PTH or PTHrP for binding can be identified using the protein and DNAs homologous to PTH DNA can be identified using fragments of the clone as probes. The sequence CC may be used for the prodn. of antibodies useful for the treatment, CC classification, prognosis and/or treatment of disorders related to CC the interaction between a cell receptor and a ligand such as in hypercalcaemia. See also R27704-16.

XX
SQ Sequence 591 AA;

Query Match 48.3%; Score 145; DB 13; Length 591;
Best Local Similarity 57.9%; Pred. No. 1.6e-10;
Matches 22; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 1 DFNHKGVAFRHCNPGTWFDFWHSLNKTKWANYSDCLREL 38
||||| :|: ||:: : |:|||||::|::|:
Db 137 dfnhkghayrdcrgswvvpghnrwtwanysecikfm 174

RESULT 10

R92277

ID R92277 standard; Protein; 591 AA.
XX AC
XX AC
XX DT
XX DT
XX DT
DE DE 18-MAY-1996 (first entry)
XX DE Rat bone PTH/PTHrP receptor.
XX Parathyroid hormone; receptor; parathormone; PTH;
KW Parathyroid hormone-related protein; PTHrP; calcium; homeostasis;
KW hypercalcaemia; hypocalcaemia; cancer.
XX OS Rattus sp.
XX XX
FH Key Location/Qualifiers
FT Region 1..192 /label= Extracellular_region
FT Region 193..211 /label= Transmembrane_region
FT Region 212..221 /label= Intracellular_region
FT Region 222..240 /label= Transmembrane_region
FT Region 241..299 /label= Extracellular_region
FT Region 300..316 /label= Transmembrane_region
FT Region 317..325 /label= Intracellular_region
FT Region 326..342 /label= Transmembrane_region
FT Region 343..364 /label= Extracellular_region
FT Region 365..383 /label= Transmembrane_region
FT Region 384..408 /label= Intracellular_region
FT Region 409..428 /label= Transmembrane_region
FT Region 429..444 /label= Intracellular_region
FT Region 445..463 /label= Transmembrane_region
FT Region 464..591 /label= Intracellular_region
XX US5494806-A.
XX XX
PD 27-FEB-1996.
XX PF
PF 05-APR-1991; 91US-0681702.


```

XX Opossum kidney PTH/PTHrP receptor prod. from clone OK-O.
DE Parathyroid hormone; related protein; calcium; antagonist;
XX antibodies; hypercalcaemia.
KW Didelphis virginiana.
XX WO9217602-A.
XX PN 15-OCT-1992.
XX PD 06-APR-1992; 92WO-US02821.
XX PF 05-APR-1991; 91US-0681702.
XX PR 06-APR-1992; 92US-0864475.
XX PA (GEO ) GEN HOSPITAL CORP OFFICE TECHNOLOGY.
XX ABou-samra A, Juppner H, Kronenberg HM, Potts JT, Schipani E;
PI Segre GV;
XX WPI; 1992-366271/44.
XX N-PSDB; Q29605.
DR - for (differential) diagnosis of hypercalcaemia, and diagnosis
XX and treatment of tumours
XX Disclosure; Fig 2; 91pp; English.
XX The parathyroid hormone/parathyroid hormone related protein (PTH/PTHrP)
CC receptor protein sequence was deduced form the DNA sequence of the
CC clone OK-O, isolated from opossum kidney (OK) cells. The clone OK-O
CC is identical to the OK-H clone except at the C-terminal tail as OK-O
CC encodes a 585 amino acid protein, OK-H encodes one of 515 amino acids.
CC The difference is attributed to a single nucleotide deleted in the OK-H
CC sequence causing a frame shift and an earlier stop codon. It is not
CC known whether OK-O and OK-H represent prods. of two separate genes or
CC are a laboratory artifact. The protein may be used in a therapeutic
CC compsn. to inhibit activation of PTH or PTHrP and thus reduce the
CC level of calcium in the blood. Cpds. capable of competing with PTH
CC or PTHrP for binding can be identified using the protein prod. and
CC DNAs homologous to PTH DNA can be identified using fragments of the
CC clone as probes. The sequence may be used for the prodn. of antibodies
CC useful for the treatment, classification, prognosis and/or treatment of
CC disorders related to the interaction between a cell receptor and a
CC ligand such as in hypercalcaemia. See also R27704-16.
XX Sequence 585 AA;
SQ
Query Match 48.7%; Score 146; DB 13; Length 585;
Best Local Similarity 57.9%; Pred. No. 1.1e-10;
Matches 22; Conservative 9; Mismatches 7; Indels 0; Gaps 0;
QY 1 DFNHKGVAFRHCNPNGTWDFMHSLNKTKWANYSDCLREL 38
Db 134 dfnhkgayrrcdsgngswelpggnrtwanysecvkfl 171
RESULT 6
R92276
ID R92276 standard; Protein; 585 AA.
XX R92276;
XX AC
XX DT
XX DE 18-MAY-1996 (first entry)
XX Opossum kidney PTH/PTHrP receptor.
XX Parathyroid hormone; receptor; parathormone; PTH;
KW Parathyroid hormone-related protein; calcium; homeostasis;
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:51:40 ; Search time 138.73 Seconds
(without alignments)
12.570 Million cell updates/sec

Title: US-09-236-468a-2_COPY_92_142

Perfect score: 300

Sequence: 1 DFNHKGVAFRHCNPGTWDF.....SDCLRFQLQPDISIGKQEFCE 51

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_36:*

- 1: /SIDSL1/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SIDSL1/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SIDSL1/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SIDSL1/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SIDSL1/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SIDSL1/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SIDSL1/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SIDSL1/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SIDSL1/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SIDSL1/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SIDSL1/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SIDSL1/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SIDSL1/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SIDSL1/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SIDSL1/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SIDSL1/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SIDSL1/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SIDSL1/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SIDSL1/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SIDSL1/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SIDSL1/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	300	100.0	541	18 W12695	G-protein parathyr
2	146	48.7	515	13 R27704	Opessum kidney PTH
3	146	48.7	515	17 R92275	Opessum kidney PTH
4	146	48.7	515	20 W73314	Parathyroid hormon
5	146	48.7	585	13 R27705	Opessum kidney PTH
6	146	48.7	585	17 R92276	Opessum kidney PTH
7	146	48.7	585	20 W73315	Parathyroid hormon
8	146	48.7	614	13 R27707	Human kidney PTH/p
9	145	48.3	591	13 R27706	Rat bone PTH/pTHr/p
10	145	48.3	591	17 R92277	Rat bone PTH/pTHr/p
11	145	48.3	591	20 W73316	Parathyroid hormon
12	145	48.3	593	17 R92278	Human kidney PTH/p

13	145	48.3	593	20 W73317	Human Parathyroid
14	76.5	25.5	477	15 R62448	Rat calcitonin rec
15	76.5	25.5	515	15 R62447	Rat calcitonin rec
16	72	24.0	472	17 R92753	Human calcitonin r
17	71	23.7	180	15 R51703	cDNA coding human
18	71	23.7	474	15 R51702	Human calcitonin r
19	71	23.7	477	15 R51704	Human placental ca
20	71	23.7	490	14 R37424	Human CTR. Homo s
21	70	23.3	553	19 W68065	Human glucagon-lik
22	69	23.0	461	17 R92220	Calcitonin gene-re
23	69	23.0	550	19 W68066	Rat glucagon-like
24	63	21.0	487	19 W55047	Alkaline phosphata
25	57.5	19.2	642	21 W77133	Marburg virus trun
26	57.5	19.2	681	21 Y70075	Marburg virus enve
27	57.5	19.2	681	21 Y77127	Marburg virus glyc
28	57.5	19.2	764	20 W94372	Human Factor B ana
29	57	19.0	1470	18 W23411	Porcine transmissi
30	56.5	18.8	136	19 W76228	Bacterial periplas
31	56.5	18.8	530	19 W55046	Human placental al
32	56.5	18.8	531	10 P91776	Germ cell alkaline
33	56.5	18.8	532	13 R20527	Human germ cell al
34	56.5	18.8	535	19 W70901	Human placental al
35	56.5	18.8	777	15 R66044	Human N-methyl-D-a
36	56.5	18.8	777	20 W85578	Human N-methyl-D-a
37	56.5	18.8	777	21 Y56115	Human clone NMDA7
38	56.5	18.8	854	15 R66052	Human NMDA recepto
39	56.5	18.8	854	20 W85586	Human N-methyl-D-a
40	56.5	18.8	854	21 Y56123	Human NMDAR1-delta
41	56.5	18.8	870	15 R66051	Human NMDA recepto
42	56.5	18.8	870	20 W85585	Human N-methyl-D-a
43	56.5	18.8	870	21 Y56122	Human NMDAR1-delta
44	56.5	18.8	875	15 R66050	Human NMDA recepto
45	56.5	18.8	875	20 W85584	N-methyl-D-asparta

ALIGNMENTS

RESULT 1

W12695 ID W12695 standard; Protein; 541 AA.

XX AC W12695;

XX DT 31-MAY-1997 (first entry)

XX DE G-protein parathyroid hormone receptor HLTG74.

XX KW G-protein parathyroid hormone receptor; HLTG74; parathormone; PTH:

XX KW calcium; signal transduction; agonist; antagonist; hypocalcaemia;

XX KW hyperphosphataemia; hypoparathyroidism; chronic tetany;

XX KW osteoporosis; hypercalcaemia; hypoparathyroidism; hypophosphataemia;

XX KW kidney stone; nephroliasis; therapy; diagnosis.

XX OS Homo sapiens.

XX PN WO9639433-A1.

XX PD 12-DEC-1996.

XX PF 05-JUN-1995; 95WO-US07085.

XX PR 05-JUN-1995; 95WO-US07085.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Li Y, Rosen CA, Ruben SM, Soppet DR;

XX DR WPI: 1997-043068/04.

XX DR N-PSDB; T59619.

XX PT Human G-protein parathyroid hormone receptor, HLTG74 - used to

PT identify (ant)agonists, used in the treatment of hypo- or

AC Q9RLW5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CALCITONIN RECEPTOR-LIKE RECEPTOR PRECURSOR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RA Tsujikawa K., Tadotsu N., Takizawa A., Hayashi T., Yamamoto H.;
RT "Mouse calcitonin receptor-like receptor.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB015595; BAA76492.1; -;
DR INTERPRO; IPR000832; -;
DR INTERPRO; IPR001688; -;
DR INTERPRO; IPR001879; -;
DR PFAM; PF00002; 7tm_2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PRINTS; PR00361; CALCITONINR.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 463 AA; 53204 MW; 338CC0B5ED2B6899 CRC64;

Query Match 35.0%; Score 53.5; DB 11; Length 463;
Best Local Similarity 47.4%; Pred. No. 3.3;
Matches 9; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

Qy 1 LQEGEG-NCFPEWDGLICW 18
Db 56 IQQAEGLYCNRRTWDGWLW 74

Search completed: November 8, 2000, 08:56:02
Job time: 513 sec

```

RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Connell M.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: U28741; AAA68325.1; -.
DR INTERPRO: IPR000561; -.
DR INTERPRO: IPR002049; -.
DR PROSITE: PS00022; EGF_1; UNKNOWN_5.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01248; LAMININ_TYPE_EGF; UNKNOWN_1.
KW Glycoprotein.
SQ SEQUENCE 264 AA; 28434 MW; C91H1CC7B4B463F4 CRC64;

Query Match 35.3%; Score 54; DB 5; Length 264;
Best Local Similarity 52.9%; Pred. No. 1.6;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 LQEGEGNCFPEWDGLIC 17
   I I I I I I I I I
DB 98 LVEGKCECFERWTGLFC 114

RESULT 14
Q9WUP2 PRELIMINARY; PRT; 463 AA.
ID Q9WUP2;
AC Q9WUP2;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CALCITONIN RECEPTOR-LIKE RECEPTOR.
CRLR GN GN
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA Derst C., Preisig-Mueller R., Gerhards J., Daut J.;
RT "Cloning and sequencing of mouse CGRP/adrenomedullin receptor
   subunits.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF146525; AAD35021.1; -.
DR INTERPRO: IPR000832; -.
DR INTERPRO: IPR001688; -.
DR INTERPRO: IPR001879; -.
DR PFAM: PF00002; 7tm_2; 1.
DR PRINTS: PR00249; GPCRSECRETIN.
DR PRINTS: PR00361; CALCITONINR.
DR PROSITE: PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE: PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 463 AA; 53234 MW; 630EC8956A58847E CRC64;

Query Match 35.0%; Score 53.5; DB 11; Length 463;
Best Local Similarity 47.4%; Pred. No. 3.3;
Matches 9; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 1 LQEGEGNCFPEWDGLICW 18
   I I I I I I I I I
DB 56 IQAEGLYCNRWDGWLW 74

RESULT 15
Q9R1W5 PRELIMINARY; PRT; 463 AA.
ID Q9R1W5;

```



```
ID O73769 PRELIMINARY; PRT; 465 AA.
AC O73769;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE PITUITARY ADENYLATE CYCLASE ACTIVATING POLYPEPTIDE TYPE 1 RECEPTOR
DE PRECURSOR.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprinidae; Cyprinidae; Carassius.
RN [1]
RP SEQUENCE FROM N.A.
RA Wong A.O.L., Leung M.Y., Shea W.L.C., Chang J.P., Chow B.K.C.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RA EMBL; AF048820; AAC15699.1; -
DR INTERPRO: IPR000832; -
DR INTERPRO: IPR001879; -
DR PFAM: PF00002; 7tm2; 1.
DR PRINTS: PR00249; GPCRSECRETIN.
DR PROSITE: PS00649; G-PROTEIN_RECEP_F2_1; 1.
DR PROSITE: PS00650; G-PROTEIN_RECEP_F2_2; 1.
SQ SEQUENCE 465 AA; 53459 MW; 9779A95EDBFDD1DC1 CRC64;

Query Match 41.8%; Score 64; DB 13; Length 465;
Best Local Similarity 47.8%; Pred. No. 0.088;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 3 EGEKCFPEWDLICWPRGTVGK 25
Db 65 DKGFCPEWDLICWPRGTVGK 87

RESULT 10
Q9V716 PRELIMINARY; PRT; 504 AA.
ID Q9V716
AC Q9V716;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CG8422 PROTEIN.
DE CG8422.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY.
RX MEDLINE; 20196006.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
```

```
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J., Ye J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003814; AAP58250.1; -
DR FLYBASE; FBgn003932; CG8422.
DR INTERPRO: IPR000832; -
DR INTERPRO: IPR001879; -
DR INTERPRO: IPR002001; -
DR INTERPRO: IPR003051; -
DR PFAM: PF00002; 7tm2; 1.
DR PRINTS: PR00249; GPCRSECRETIN.
DR PRINTS: PR01127; DIUHORMONER.
DR PRINTS: PR01279; CRFRECEPTOR.
SQ SEQUENCE 504 AA; 56343 MW; 007422A7D1FA2B64 CRC64;

Query Match 39.9%; Score 61; DB 5; Length 504;
Best Local Similarity 42.1%; Pred. No. 0.27;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 6 GNCFPEDGLICWPRGTVG 24
Db 51 GHCLTQFDSILCWPRGTVG 69

RESULT 11
O73768 PRELIMINARY; PRT; 438 AA.
ID O73768
AC O73768;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE GROWTH-HORMONE RELEASING HORMONE-LIKE PEPTIDE RECEPTOR.
DE OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprinidae; Cyprinidae; Carassius.
RN [1]
RP SEQUENCE FROM N.A.
RA Chan K.W., Yu K.L., Rivier J., Chow B.K.C.;
RL Neuroendocrinology 0:0-0(1998).
DR EMBL; AF048819; AAC15698.1; -
DR INTERPRO: IPR000832; -
DR INTERPRO: IPR001879; -
DR PFAM: PF00002; 7tm2; 1.
DR PRINTS: PR00249; GPCRSECRETIN.
DR PROSITE: PS00650; G-PROTEIN_RECEP_F2_2; UNKNOWN_1.
SQ SEQUENCE 438 AA; 50305 MW; C3516A52D692C339 CRC64;

Query Match 39.2%; Score 60; DB 13; Length 438;
Best Local Similarity 33.3%; Pred. No. 0.33;
Matches 8; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

Qy 2 QEGEGNCFPEWDLICWPRGTVGK 25
Db 55 QSQQTGCWTDMDGRCWOTAKTQ 78

RESULT 12
```

```
FT NON_TER 1 1
FT NON_TER 41 41
SQ SEQUENCE 41 AA; 4361 MW; AA2CEAF541B52EFA CRC64;

Query Match 41.8%; Score 64; DB 6; Length 41;
Best Local Similarity 41.9%; Pred. No. 0.0073;
Matches 13; Conservative 2; Mismatches 10; Indels 6; Gaps 1;

Qy 1 LOEGEG-----NCFPEWDGLICPRGTGK 25
Db 6 LQAEGMPNSTLGCPRWDGLLCWPTAGSGE 36

RESULT 6
Q9TUJ1
ID Q9TUJ1 PRELIMINARY; PRT; 404 AA.
AC Q9TUJ1;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE GROWTH HORMONE-RELEASING HORMONE RECEPTOR SHORT FORM.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A.
RA Takata M., Tarumi O., Watanabe S., Sekikawa K.;
RT "Molecular cloning of bovine growth hormone-releasing hormone receptor
RT cDNA.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022596; BA84959.1; -.
DR INTERPRO; IPR000832; -.
DR INTERPRO; IPR001771; -.
DR INTERPRO; IPR001879; -.
DR PFAM; PF00002; 7tm.2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 404 AA; 44868 MW; D394FB43BECAB4AC CRC64;

Query Match 41.8%; Score 64; DB 6; Length 404;
Best Local Similarity 41.9%; Pred. No. 0.076;
Matches 13; Conservative 2; Mismatches 10; Indels 6; Gaps 1;

Qy 1 LOEGEG-----NCFPEWDGLICPRGTGK 25
Db 42 LQAEGMPNSTLGCPRWDGLLCWPTAGSGE 72

RESULT 7
Q9WU99
ID Q9WU99 PRELIMINARY; PRT; 439 AA.
AC Q9WU99;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE GHRH RECEPTOR BETA GHRHR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-SPRAGUE-DAWLEY; TISSUE-PITUITARY;
RX MEDLINE; 99061817.
RA Zeitler P., Stevens P., Siriwardana G.;
RT "Functional GHRH receptor carboxyl terminal isoforms in normal and
RT dwarf (dw) rats.";
RL J. Mol. Endocrinol. 21:363-371(1998).
```

```
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-PITUITARY;
RA Zeitler P., Stevens P., Siriwardana G.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF122055; AAD26335.1; -.
DR INTERPRO; IPR000832; -.
DR INTERPRO; IPR001771; -.
DR INTERPRO; IPR001879; -.
DR PFAM; PF00002; 7tm.2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PROSITE; PR01154; VIPRECEPTOR.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 439 AA; 48763 MW; F1BB2C9A855D24ED CRC64;

Query Match 41.8%; Score 64; DB 11; Length 439;
Best Local Similarity 48.0%; Pred. No. 0.083;
Matches 12; Conservative 1; Mismatches 6; Indels 6; Gaps 1;

Qy 1 LOEGEG-----NCFPEWDGLICWP 19
Db 42 LQAEGTNNSSMGCPGTWDLGLLWCP 66

RESULT 8
Q9TUJ0
ID Q9TUJ0 PRELIMINARY; PRT; 441 AA.
AC Q9TUJ0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE GROWTH HORMONE-RELEASING HORMONE RECEPTOR LONG FORM.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A.
RA Masuhiro T., Tarumi O., Watanabe S., Sekikawa K.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Takata M., Tarumi O., Watanabe S., Sekikawa K.;
RT "Molecular cloning of bovine growth hormone-releasing hormone receptor
RT cDNA.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022597; BA84960.1; -.
DR INTERPRO; IPR000832; -.
DR INTERPRO; IPR001771; -.
DR INTERPRO; IPR001879; -.
DR PFAM; PF00002; 7tm.2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PROSITE; PR01154; VIPRECEPTOR.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 441 AA; 49251 MW; B8E11893EFC3EBB8 CRC64;

Query Match 41.8%; Score 64; DB 6; Length 441;
Best Local Similarity 41.9%; Pred. No. 0.084;
Matches 13; Conservative 2; Mismatches 10; Indels 6; Gaps 1;

Qy 1 LOEGEG-----NCFPEWDGLICWPRTGK 25
Db 42 LQAEGMPNSTLGCPRWDGLLCWPTAGSGE 72

RESULT 9
O73769
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:56:01 ; Search time 152.43 Seconds
(without alignments)
15,314 Million cell updates/sec

Title: US-09-236-468a-2_COPY_56_80
Perfect score: 153
Sequence: 1 LQEGGNCPEWDGLICWPGTGVK 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 297973

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_14.*
1: sp-archaea.*
2: sp-bacteria.*
3: sp-fungi.*
4: sp-human.*
5: sp-invertebrate.*
6: sp-mammal.*
7: sp-mhc.*
8: sp-organelle.*
9: sp-phase.*
10: sp-plant.*
11: sp-rodent.*
12: sp-virus.*
13: sp-vertebrate.*
14: sp-unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102	66.7	536	13 Q9PVD3	Q9pvd3 brachydanio
2	91	59.5	542	13 Q9PVD2	Q9pvd2 brachydanio
3	90	58.8	575	13 Q9PWB7	Q9pwb7 brachydanio
4	76	49.7	595	6 Q9TU31	Q9tu31 canis famill
5	64	41.8	41	6 Q77678	Q77678 bos taurus
6	64	41.8	404	6 Q9TUJ1	Q9tuj1 bos taurus
7	64	41.8	439	11 Q9WU99	Q9wu99 rattus norv
8	64	41.8	441	6 Q9TUJ0	Q9tuj0 bos taurus
9	64	41.8	465	13 Q73769	Q73769 carassius a
10	61	39.9	504	5 Q9V716	Q9v716 drosophila
11	60	39.2	438	13 Q73768	Q73768 carassius a
12	58	37.9	444	13 Q9YHC6	Q9yhc6 rana ridibu
13	54	35.3	264	5 Q20043	Q20043 caenorhabdi
14	53.5	35.0	463	11 Q9WUP2	Q9wup2 mus musculu
15	53.5	35.0	463	11 Q9RIW5	Q9riw5 mus musculu
16	53.5	35.0	463	11 Q9QXH8	Q9qxh8 mus musculu
17	53	34.6	1381	11 P97846	P97846 rattus norv
18	53	34.6	1384	4 P78357	P78357 homo sapien
19	53	34.6	1385	11 O54991	O54991 mus musculu

20 52 34.0 444 5 Q9V6N5
21 52 34.0 459 11 Q9RI78
22 51.5 33.7 183 5 O01522
23 51.5 33.7 561 5 Q9VYH9
24 51 33.3 116 5 Q9XZE7
25 51 33.3 465 13 Q9PTK1
26 50.5 33.0 1645 5 Q9U263
27 50.5 33.0 1655 5 O44498
28 50 32.7 149 5 Q22745
29 50 32.7 311 2 Q9X7G5
30 50 32.7 388 5 Q9V6C7
31 50 32.7 403 5 O18375
32 50 32.7 559 5 Q9VZ44
33 50 32.7 633 5 Q9W4Y2
34 48 31.4 220 11 Q63404
35 48 31.4 721 13 Q91902
36 48 31.4 1574 11 O88281
37 47.5 31.0 452 13 Q9PSZ3
38 47 30.7 1530 6 Q28146
39 46.5 30.4 332 6 O18949
40 46.5 30.4 383 6 Q29411
41 46 30.1 147 2 P75077
42 46 30.1 211 5 O44443
43 46 30.1 266 2 Q9ZA31
44 46 30.1 460 3 Q92402
45 46 30.1 777 11 Q64548

ALIGNMENTS

RESULT 1
Q9PVD3
ID Q9PVD3 PRELIMINARY; PRT; 536 AA.
AC Q9PVD3;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE PARATHYROID HORMONE RECEPTOR PTHIR.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprinidae; Rasbora; Danio.
RN [1]
RP SEQUENCE FROM N.A.
RA Rubin D.A., Jueppner H.;
RT "Zebrafish Express the Common Parathyroid Hormone/Parathyroid Hormone-
RT related Peptide Receptor (PTHIR) and a Novel Receptor (PTH3R) That Is
RT Preferentially Activated by Mammalian and Fuguish Parathyroid
RT Hormone-related Peptide."
RL J. Biol. Chem. 274:28185-28190(1999).
DR EMBL: AF132084; AAF01265.1; -.
DR INTERPRO: IPR000832; -.
DR INTERPRO: IPR001879; -.
DR INTERPRO: IPR002170; -.
DR PFAM: PF00002; 7tm.2; 1.
DR PRINTS: PR00249; GPCRSECRETIN.
DR PRINTS: PR00393; PTHORMONER.
DR PRINTS: PR00393; PTHORMONER.
DR PROSITE: PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE: PS00650; G_PROTEIN_RECEP_F2_2; 1.
DR PROSITE: PS00650; G_PROTEIN_RECEP_F2_2; 1.
DR PROSITE: PS00650; G_PROTEIN_RECEP_F2_2; 1.
SQ SEQUENCE 536 AA; 61454 MW; 653B0A56DB2C8FA9 CRC64;

Query Match 66.7%; Score 102; DB 13; Length 536;
Best Local Similarity 66.7%; Pred. No. 2.1e-07;
Matches 14; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 EGNCFPEWDGLICWPGTGVK 25

DB 60 EGSCLPPEWDGLICWPGVPGK 80

THIS PAGE BLANK (USPTO)

```
FT SIGNAL          1 31 POTENTIAL.
FT CHAIN          32 458 VASOACTIVE INTESTINAL POLYPEPTIDE
FT RECEPTOR 1.
FT DOMAIN         32 143 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM      144 168 1 (POTENTIAL)
FT DOMAIN         169 175 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM      176 195 2 (POTENTIAL).
FT DOMAIN         196 217 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM      218 241 3 (POTENTIAL).
FT DOMAIN         242 255 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM      256 277 4 (POTENTIAL).
FT DOMAIN         278 293 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM      294 317 5 (POTENTIAL).
FT DOMAIN         318 342 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM      343 362 6 (POTENTIAL).
FT DOMAIN         363 374 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM      375 394 7 (POTENTIAL).
FT DOMAIN         395 458 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD      59 59 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD      70 70 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD     101 101 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD     105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE      458 AA; 51479 MW; E166E4D6B3BE1189 CRC64;
```

Query Match 36.9%; Score 56.5; DB 1; Length 458;
Best Local Similarity 42.3%; Pred. No. 0.77;
Matches 11; Conservative 1; Mismatches 11; Indels 3; Gaps 1;

QY 1 LOEGEGNCFPEDWGLICW---PRGTV 23
|: | | | | | | | | | |
DB 57 LENETSGCGSKMWDNLTCWPATPRGQV 82

Search completed: November 8, 2000, 09:03:44
Job time: 854 sec

```

FT DOMAIN 281 296 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 297 320 5 (POTENTIAL).
FT DOMAIN 321 343 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 344 361 6 (POTENTIAL).
FT DOMAIN 362 373 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 374 395 7 (POTENTIAL).
FT DOMAIN 396 474 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPIC 384 397 MISSING (IN ISOFORM CTDELTAEL13).
SQ SEQUENCE 474 AA; 55235 MW; 85A631CAE1C61412 CRC64;

Query Match 38.6%; Score 59; DB 1; Length 474;
Best Local Similarity 46.2%; Pred. No. 0.35;
Matches 12; Conservative 3; Mismatches 7; Indels 4; Gaps 2;

Qy 3 ELEG-NCFPNDGLICW---PRGTVG 24
:||||| ||||| ||||| |||||
Db 66 QGEGPCNRTWDGMCWDDTPAGVLG 91

RESULT 14
DIHR_ACHDO
ID DIHR_ACHDO STANDARD; PRT; 441 AA.
AC Q16983;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE DIURETIC HORMONE RECEPTOR PRECURSOR (DH-R).
OS Acheta domestica (House cricket).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Gryllota; Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae;
OC Gryllinae; Acheta.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-MALPIGHIAN TUBULES;
RA MEDLINE; 96189577.
RX Reagan J.D.;
RT "Molecular cloning and function expression of a diuretic hormone
receptor from the house cricket, Acheta domestica."
RL Insect Biochem. Mol. Biol. 26:1-6(1996).
CC -!- FUNCTION: RECEPTOR FOR THE INSECT DIURETIC HORMONE. THE ACTIVITY
OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL
CYCLASE.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MALPIGHIAN TUBULES.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U15959; AAC47000.1; -
GCRDB; GCR_1380; -
DR INTERPRO; IPR000832; -
DR INTERPRO; IPR002001; -
DR PFAM; PF00002; 7tm2.1;
DR PRINTS; PR00249; GPCRSECRETIN.
DR PRINTS; PR01127; DIUHORMONER.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2.1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2.2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 ?
FT CHAIN ? 441 DIURETIC HORMONE RECEPTOR.
FT DOMAIN ? 134 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 135 158 1 (POTENTIAL).

```

```

FT DOMAIN 159 166 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 167 187 2 (POTENTIAL).
FT DOMAIN 188 194 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 195 224 3 (POTENTIAL).
FT DOMAIN 225 238 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 239 260 4 (POTENTIAL).
FT DOMAIN 261 291 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 292 315 5 (POTENTIAL).
FT DOMAIN 316 338 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 339 357 6 (POTENTIAL).
FT DOMAIN 358 371 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 372 391 7 (POTENTIAL).
FT DOMAIN 392 441 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 107 107 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 441 AA; 49538 MW; 501915AC2E776C5C CRC64;

Query Match 36.9%; Score 56.5; DB 1; Length 441;
Best Local Similarity 33.3%; Pred. No. 0.74;
Matches 11; Conservative 2; Mismatches 11; Indels 9; Gaps 1;

Qy 1 LQEGEG-----NCFPNDGLICWPRGTVG 24
| ||| | ||| | ||| |
Db 46 LMEGDAPDAAPDAPLRCAIAWDGVNCWPETPAG 78

RESULT 15
VIPR_PIG
ID VIPR_PIG STANDARD; PRT; 458 AA.
AC Q28992;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE VASOACTIVE INTESTINAL POLYPEPTIDE RECEPTOR I PRECURSOR (VIP-R-1)
DE (PITUITARY ADENYLATE CYCLASE ACTIVATING POLYPEPTIDE TYPE II RECEPTOR)
DE (PACAP TYPE II RECEPTOR) (PACAP-R-2).
GN VIPR1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RA Hsiung H.M., Smith D.P., Hyslop P.A., Heiman M.L., Hassan H.A.,
RA Zhang X.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS IS A RECEPTOR FOR VIP. THE ACTIVITY OF THIS
RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL
CYCLASE (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U49434; AAA93390.1; -
GCRDB; GCR_1565; -
DR INTERPRO; IPR000832; -
DR INTERPRO; IPR001571; -
DR INTERPRO; IPR001771; -
DR PFAM; PF00002; 7tm2.1;
DR PRINTS; PR00249; GPCRSECRETIN.
DR PRINTS; PR00491; VASOACTIVEIPR.
DR PRINTS; PR01154; VIP1RECEPTOR.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2.1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2.2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.

```

AC Q63118;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE CALCITONIN GENE-RELATED PEPTIDE TYPE 1 RECEPTOR PRECURSOR (CGRP TYPE 1
 DE RECEPTOR).
 GN CGRPR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WISTAR; TISSUE-LUNG;
 RX MEDLINE; 94037821.
 RA Njuki F., Nicholl C.G., Howard A., Mak J.C., Barnes P.J., Girgis S.I.,
 RA Legon S.;
 RT "A new calcitonin-receptor-like sequence in rat pulmonary blood
 RT vessels.";
 RL Clin. Sci. 85:385-388(1993).
 CC -!- FUNCTION: THIS IS A RECEPTOR FOR CALCITONIN GENE-RELATED PEPTIDE
 CC TYPE 1. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS
 CC WHICH ACTIVATE ADENYLYL CYCLASE (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X70658; CAA49997.1; -;
 DR GCRDB; GCR_0951; -;
 DR INTERPRO; IPR000832; -;
 DR PFAM; PF00002; 7tm_2; 1.
 DR PRINTS; PR00249; GPCRSECRETIN.
 DR PROSITE; PS00649; G_PROTEIN_RECP_F2_1; 1.
 DR PROSITE; PS00650; G_PROTEIN_RECP_F2_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 464
 FT CALCITONIN GENE-RELATED PEPTIDE TYPE 1
 FT RECEPTOR.
 FT DOMAIN 24 146
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 147 166
 FT 1 (POTENTIAL).
 FT DOMAIN 167 173
 FT CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 174 193
 FT 2 (POTENTIAL).
 FT DOMAIN 194 213
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 214 236
 FT 3 (POTENTIAL).
 FT DOMAIN 237 253
 FT CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 254 273
 FT 4 (POTENTIAL).
 FT DOMAIN 274 289
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 290 313
 FT 5 (POTENTIAL).
 FT DOMAIN 314 336
 FT CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 337 354
 FT 6 (POTENTIAL).
 FT DOMAIN 355 366
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 367 388
 FT 7 (POTENTIAL).
 FT DOMAIN 389 464
 FT CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 30 30
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 66 66
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 118 118
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 123 123
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 128 128
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 129 129
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 464 AA; 53329 MW; 7373EC94B0FBC33C CRC64;

Query Match 38.9%; Score 59.5; DB 1; Length 464;
 Best Local Similarity 52.6%; Pred. No. 0.29;
 Matches 10; Conservativity 3; Mismatches 5; Indels 1; Gaps 1;

QY 1 LOEGEG-NCPPWDGLICW 18

DB 57 IQOGEGLYCNRTWDGLCW 75
 RESULT 13
 CALR_RABIT
 ID CALR_RABIT STANDARD; PRT; 474 AA.
 AC P79222; P79223;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DE CALCITONIN RECEPTOR PRECURSOR (CT-R).
 GN CALCR.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NEW ZEALAND WHITE;
 RX MEDLINE; 97094876.
 RA Shyu J.F., Inoue D., Baron R., Horne W.C.;
 RT "The deletion of 14 amino acids in the seventh transmembrane domain
 RT of a naturally occurring calcitonin receptor isoform alters ligand
 RT binding and selectively abolishes coupling to phospholipase C.";
 RL J. Biol. Chem. 271:31127-31134(1996).
 CC -!- FUNCTION: THIS IS A RECEPTOR FOR CALCITONIN. THE ACTIVITY OF
 CC THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL
 CC CYCLASE. THE CALCITONIN RECEPTOR IS THOUGHT TO COUPLE TO THE
 CC HETEROTRIMERIC GUANOSINE TRIPHOSPHATE-BINDING PROTEIN THAT IS
 CC SENSITIVE TO CHOLERA TOXIN.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- ALTERNATIVE PRODUCTS: TWO FORMS ARE PRODUCED BY ALTERNATIVE
 CC SPLICING. THE SHORT FORM, KNOWN AS CTDELTA13, LACKS 14 RESIDUES
 CC IN TM7 AND HAS ALTERED LIGAND BINDING AND ABOLISHED COUPLING TO
 CC PHOSPHOLIPASE C.
 CC -!- TISSUE SPECIFICITY: THE TWO ISOFORMS ARE EXPRESSED IN A TISSUE-
 CC SPECIFIC MANNER, WITH CTDELTA13 ACCOUNTING FOR LESS THAN 15% OF
 CC THE TOTAL CALCITONIN RECEPTOR MRNA IN OSTEOCLASTS, KIDNEY, AND
 CC BRAIN, BUT COMPRISING AT LEAST 50% OF THE TRANSCRIPTS IN SKELETAL
 CC MUSCLE AND LUNG.
 CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U66365; AAC48687.1; -;
 DR EMBL; U73126; AAB38258.1; -;
 DR GCRDB; GCR_1570; -;
 DR GCRDB; GCR_1571; -;
 DR INTERPRO; IPR000832; -;
 DR INTERPRO; IPR001688; -;
 DR PFAM; PF00002; 7tm_2; 1.
 DR PRINTS; PR00249; GPCRSECRETIN.
 DR PRINTS; PR00361; CALCITONINR.
 DR PROSITE; PS00649; G_PROTEIN_RECP_F2_1; 1.
 DR PROSITE; PS00650; G_PROTEIN_RECP_F2_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 Alternative splicing
 FT SIGNAL 1 22
 FT CHAIN 23 474
 FT CALCITONIN RECEPTOR.
 FT DOMAIN 23 153
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 154 173
 FT 1 (POTENTIAL).
 FT DOMAIN 174 180
 FT CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 181 200
 FT 2 (POTENTIAL).
 FT DOMAIN 201 220
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 221 243
 FT 3 (POTENTIAL).
 FT DOMAIN 244 260
 FT CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 261 280
 FT 4 (POTENTIAL).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/annouche/> or send an email to license@isb-sib.ch).

```
CC EMBL; L01406; AAA35890.1; -.
CC EMBL; L09237; AAA58619.1; -.
CC EMBL; L09237; AAA58619.1; -.
CC EMBL; U42225; AAB37758.1; -.
CC EMBL; U42222; AAB37758.1; JOINED.
CC EMBL; U42223; AAB37758.1; JOINED.
CC EMBL; U42224; AAB37758.1; JOINED.
CC PIR; A45363; A45363.
CC GCRDB; GCR_0461; -.
CC GCRDB; GCR_0516; -.
CC GCRDB; GCR_2097; -.
CC MIM; 139191; -.
CC INTERPRO; IPR000832; -.
CC PFAM; PF00002; 7tm_2; 1.
CC PRINTS; PR00249; GPCRSECRETIN.
CC PROSITE; PS00649; G_PROTEIN_RECEP_F2.1; 1.
CC PROSITE; PS00650; G_PROTEIN_RECEP_F2.2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 22
FT CHAIN 23 423
FT GROWTH HORMONE-RELEASING HORMONE
FT RECEPTOR.
FT DOMAIN 23 132
FT TRANSMEM 133 152
FT DOMAIN 153 162
FT TRANSMEM 163 181
FT DOMAIN 182 207
FT TRANSMEM 205 224
FT DOMAIN 228 240
FT TRANSMEM 241 262
FT DOMAIN 263 280
FT TRANSMEM 281 304
FT DOMAIN 305 329
FT TRANSMEM 330 348
FT DOMAIN 349 361
FT TRANSMEM 362 381
FT DOMAIN 382 423
FT CARBOHYD 50 50
FT CONFLICT 178 178
FT CONFLICT A -> R (IN REF. 1).
SQ SEQUENCE 423 AA; 47402 MW; C9C5E2E7D6649E06 CRC64;

Query Match 39.9%; Score 61; DB 1; Length 423;
Best Local Similarity 50.0%; Pred. No. 0.16;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Oy 8 CFPEDWGLICWPRGTGK 25
Db 55 CPATWDGLLCWPTAGSGE 72
| | | | | | | | | |
| | | | | | | | | |

RESULT 11
GRFR_PIG STANDARD; PRT; 423 AA.
AC P34999; Q28993;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE GROWTH HORMONE-RELEASING HORMONE RECEPTOR PRECURSOR (GHRH RECEPTOR)
DE (GRF RECEPTOR) (GRFR).
GN GHRHR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PITUITARY;
```

```
RX MEDLINE; 94020094.
RA Hsiung H.M., Smith D.P., Zhang X.-Y., Bennett T., Rostock P.R. Jr.,
RA Lai M.-H.;
RT "Structure and functional expression of a complementary DNA for
RT porcine growth hormone-releasing hormone receptor.";
RL Neuropeptides 25:1-10(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Hazem H.A., Zhang X., Smith D.P., Heiman M.L., Hsiung H.M.;
RA Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS IS A RECEPTOR FOR GRF. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL
CC CYCLASE.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: PITUITARY GLAND. IS ALSO DETECTED IN THE
CC LYMPHOCYTES AND THYMOCYTES.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/annouche/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL; L11869; AAA31047.1; -.
CC EMBL; U49435; AAA93391.1; -.
CC GCRDB; GCR_0650; -.
CC GCRDB; GCR_1560; -.
CC INTERPRO; IPR000832; -.
CC PFAM; PF00002; 7tm_2; 1.
CC PRINTS; PR00249; GPCRSECRETIN.
CC PROSITE; PS00649; G_PROTEIN_RECEP_F2.1; 1.
CC PROSITE; PS00650; G_PROTEIN_RECEP_F2.2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
CC Alternative splicing.
FT SIGNAL 1 22
FT CHAIN 23 423
FT GROWTH HORMONE-RELEASING HORMONE
FT RECEPTOR.
FT DOMAIN 23 130
FT TRANSMEM 131 154
FT DOMAIN 155 161
FT TRANSMEM 162 174
FT DOMAIN 175 202
FT TRANSMEM 203 227
FT DOMAIN 228 240
FT TRANSMEM 241 262
FT DOMAIN 263 280
FT TRANSMEM 281 304
FT DOMAIN 305 329
FT TRANSMEM 330 348
FT DOMAIN 349 361
FT TRANSMEM 362 381
FT DOMAIN 382 423
FT CARBOHYD 50 50
FT CONFLICT 419 423
FT CONFLICT T (IN REF. 1).
SQ SEQUENCE 423 AA; 47199 MW; 6768357DCB4AD603 CRC64;

Query Match 39.9%; Score 61; DB 1; Length 423;
Best Local Similarity 36.0%; Pred. No. 0.16;
Matches 9; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

Oy 1 LOEGEGNCFPEDWGLICWPRGTGK 25
Db 48 MANSSSGCPRTWDGLLCWPTAGSGE 72
| | | | | | | | | |
| | | | | | | | | |

RESULT 12
CGRR_RAT STANDARD; PRT; 464 AA.
ID CGRR_RAT
```

FT	DOMAIN	23	132	RECEPTOR.
FT	TRANSSEM	133	152	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	153	162	1 (POTENTIAL).
FT	TRANSSEM	163	181	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	182	204	2 (POTENTIAL).
FT	TRANSSEM	205	227	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	228	240	3 (POTENTIAL).
FT	TRANSSEM	241	262	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	263	280	4 (POTENTIAL).
FT	TRANSSEM	281	304	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	305	370	5 (POTENTIAL).
FT	TRANSSEM	371	389	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	390	402	6 (POTENTIAL).
FT	TRANSSEM	403	422	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	423	464	7 (POTENTIAL).
FT	CARBOHYD	50	50	CYTOPLASMIC (POTENTIAL).
FT	VARSPLIC	325	365	N-LINKED (GLUCNAC. . .) (POTENTIAL).
FT	CONFLICT	233	233	MISSING (IN SHORT ISOFORM).
FT	CONFLICT	305	305	S -> W (IN REF. 2).
FT	CONFLICT	305	305	C -> S (IN REF. 2).
ST	SEQUENCE	464 AA;	51521 MW;	5FF49A7D79A4114 CRC64;

Query Match 41.8%; Score 64; DB 1; Length 464;
 Best Local Similarity 48.0%; Pred. No. 0.067;
 Matches 12; Conservative 1; Mismatches 6; Indels 6; Gap

QY	1	LOEGEG-----NCFPEWDGLICWP	19
DB	42	LOAAGTNNSSMGCGPTWDGLLCWP	66

RESULT 10

GRPR_HUMAN	STANDARD;	PRT;	423 AA.
ID	GRPR_HUMAN	Q02643; Q99863;	
AC	01-OCT-1993 (Rel. 27, Created)		
DT	15-JUL-1998 (Rel. 36, Last sequence update)		
DT	15-JUL-1998 (Rel. 36, Last annotation update)		
DE	GROWTH HORMONE-RELEASING HORMONE RECEPTOR PRECURSOR (GHRH RECEPTOR)		
DE	(GRF RECEPTOR) (GRFR).		
GN	GHRHR.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
	[1]		
RN	SEQUENCE FROM N.A.		
RC	TISSUE=PITUITARY;		
RC	MEDLINE; 93078807.		
RA	MAYO K.E.;		
RT	"Molecular cloning and expression of a pituitary-specific receptor		
RT	for growth hormone-releasing hormone.";		
RL	Mol. Endocrinol. 6:1734-1744(1992).		
RN	[2]		
RN	SEQUENCE FROM N.A.		
RC	TISSUE=PITUITARY;		
RC	MEDLINE; 93188867.		
RX	Gyllian B.D.; Harrison J.K.; Zysk J.R., Lyons C.E. Jr., Lynch K.R.,		
RA	Thornier M.O.;		
RT	"Molecular cloning and expression of a human anterior pituitary		
RT	receptor for growth hormone-releasing hormone.";		
RL	Mol. Endocrinol. 7:77-84(1993).		
RN	[3]		
RN	SEQUENCE OF 37-271 FROM N.A.		
RP	Tang J., Collu R.;		
RL	submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.		
CC	-1- RECEPTION: THIS IS A RECEPTOR FOR GRF. THE ACTIVITY OF THIS		
CC	RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYL		
CC	CYCLASE.		
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.		
CC	-1- TISSUE SPECIFICITY: PITUITARY GLAND.		
CC	-1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.		


```

Qy 8 CFPWDGLICWPRGTGVK 25
Db 108 CLPWNIVCWPLGAPGE 125

RESULT 5
PTRR_RAT STANDARD: PRT; 591 AA.
AC P25961;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR
DE PRECURSOR (PTH/PTHr RECEPTOR).
GN PTHr1 OR PTHr.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1]
RC SEQUENCE FROM N.A.
RC TISSUE-BONE;
RX MEDLINE: 92212903.
RA Abou-Samra A.-B., Jueppner H., Force T., Freeman M.W., Kong X.-F.,
RA Schipani E., Urena P., Richards J., Bonventre J.V., Potts J.F. Jr.,
RA Kronenberg H.M., Segre G.V.;
RT "Expression cloning of a common receptor for parathyroid hormone and
RT parathyroid hormone-related peptide from rat osteoblast-like cells: a
RT single receptor stimulates intracellular accumulation of both cAMP
RT and inositol trisphosphates and increases intracellular free
RT calcium."
RL Proc. Natl. Acad. Sci. U.S.A. 89:2732-2736(1992).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE: 94292182.
RA Pausova Z., Bourdon J., Clayton D., Mattei M.-G., Seldin M.F.,
RA Janicic N., Riviere M., Szpirer J., Levan G., Szpirer C.;
RT "Cloning of a parathyroid hormone/parathyroid hormone-related peptide
RT receptor (PTHr) cDNA from a rat osteosarcoma (UMR 106) cell line:
RT chromosomal assignment of the gene in the human, mouse, and rat
RT genomes."
RL Genomics 20:20-26(1994).
CC -!- FUNCTION: THIS IS A RECEPTOR FOR PARATHYROID HORMONE AND FOR
CC PARATHYROID HORMONE-RELATED PEPTIDE. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATES ADENYLYL
CC CYCLASE AND ALSO A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER
CC SYSTEM.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M7184; AAA41811.1; -.
CC DR EMBL: L19475; AAA68098.1; -.
CC DR GCRDB: GCR_0206; -.
CC DR GCRDB: GCR_0938; -.
CC DR INTERPRO: IPR000832; -.
CC DR INTERPRO: IPR002170; -.
CC DR PFAM: PF00002; 7tm.2; 1.
CC DR PRINTS: PR00249; GPCRSECRETIN.
CC DR PRINTS: PR00393; PTHORMONER.
CC DR PROSITE: PS00649; G_PROTEIN_RECEP_F2_1; 1.
CC DR PROSITE: PS00650; G_PROTEIN_RECEP_F2_2; 1.
CC DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
CC SIGNAL 1 26
CC FT CHAIN 27 591
CC PARATHYROID HORMONE/PARATHYROID HORMONE-
CC RELATED PEPTIDE RECEPTOR.

```

```

FT DOMAIN 27 188 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 189 212 1 (POTENTIAL).
FT DOMAIN 214 219 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 220 239 2 (POTENTIAL).
FT DOMAIN 240 282 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 283 306 3 (POTENTIAL).
FT DOMAIN 307 320 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 321 342 4 (POTENTIAL).
FT DOMAIN 343 361 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 362 382 5 (POTENTIAL).
FT DOMAIN 383 409 6 (POTENTIAL).
FT TRANSMEM 410 428 7 (POTENTIAL).
FT DOMAIN 429 440 8 (POTENTIAL).
FT TRANSMEM 441 463 9 (POTENTIAL).
FT DOMAIN 464 591 10 (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 161 161 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 176 176 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 591 AA; 66260 MW; 21944F3051B9E9C1 CRC64;

Query Match 51.6%; Score 79; DB 1; Length 591;
Best Local Similarity 55.6%; Pred. No. 0.00063;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 8 CFPWDGLICWPRGTGVK 25
Db 108 CLPWNIVCWPLGAPGE 125

RESULT 6
PTRR_PIG STANDARD: PRT; 585 AA.
AC P50133;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR
DE PRECURSOR (PTH/PTHr RECEPTOR).
GN PTHr1 OR PTHr.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
[1]
RX MEDLINE: 96305358.
RA Black E.C., Smith D.P., Zhang X.Y., Frolik C.A., Harvey A.,
RA Chandrasekhar S., Hsiung H.M.;
RT "Structure and functional expression of a complementary DNA for
RT porcine parathyroid hormone/parathyroid hormone-related peptide
RT receptor."
RL Blochum Biophys. Acta 1307:339-347(1996).
CC -!- FUNCTION: THIS IS A RECEPTOR FOR PARATHYROID HORMONE AND FOR
CC PARATHYROID HORMONE-RELATED PEPTIDE. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATES ADENYLYL
CC CYCLASE AND ALSO A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER
CC SYSTEM (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U18315; AAC48619.1; -.
CC DR GCRDB: GCR_1607; -.
CC DR INTERPRO: IPR000832; -.
CC DR INTERPRO: IPR002170; -.

```

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

DR EMBL; M74445; AAA30979.1; -;
 DR PIR; A39286; A39286.
 DR GCRDB; GCR_0204; -;
 DR INTERPRO; IPR000832; -;
 DR INTERPRO; IPR002170; -;
 DR PFAM; PF000002; 7tm_2; 1.
 DR PRINTS; PR00249; GPCRSECRETIN.
 DR PRINTS; PR00393; PTRHORMONER.
 DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
 DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 585
 FT DOMAIN 27 185
 FT TRANSMEM 186 209
 FT DOMAIN 210 216
 FT TRANSMEM 217 236
 FT DOMAIN 237 276
 FT TRANSMEM 277 300
 FT DOMAIN 301 314
 FT TRANSMEM 315 336
 FT DOMAIN 337 355
 FT TRANSMEM 356 376
 FT DOMAIN 377 403
 FT TRANSMEM 404 422
 FT DOMAIN 423 434
 FT TRANSMEM 435 457
 FT DOMAIN 458 585
 FT CARBOHYD 148 148
 FT CARBOHYD 158 158
 FT CARBOHYD 163 163
 FT CARBOHYD 173 173
 SQ SEQUENCE 585 AA; 65963 MW; 34900384CD6DF477 CRC64;

Query Match 58.2%; Score 89; DB 1; Length 585;
 Best Local Similarity 57.1%; Pred. No. 2.4e-05;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 5 EGNCFPEWGLICWPRTGVGK 25
 :| | | | | :| | | | |
 Db 102 DGFCLPEWDNIVGWPGVPGK 122

RESULT 4
 PTHR_MOUSE STANDARD; PRT; 591 AA.
 AC P41593; Q62119;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR
 DE PRECURSOR (PTH/PTHR RECEPTOR).
 GN PTHRI OR PTHR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/HEHA;
 RX MEDLINE; 95034305.
 RA Karperien M., van Dijk T.B., Hoeijmakers T., Cremers F.,
 RA Abou-Samra A.-B., Boonstra J., de Laat S.W., Delize L.H.K.;
 RT "Expression pattern of parathyroid hormone/parathyroid hormone
 RT related peptide receptor mRNA in mouse postimplantation embryos
 RT indicates involvement in multiple developmental processes.";

RL Mech. Dev. 47:29-42(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C;
 RX MEDLINE; 94255468.
 RA McCuaig K.A., Clarke J.C., White J.H.;
 RT "Molecular cloning of the gene encoding the mouse parathyroid
 RT hormone/parathyroid hormone-related peptide receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:5051-5055(1994).
 CC -!- FUNCTION: THIS IS A RECEPTOR FOR PARATHYROID HORMONE AND FOR
 CC PARATHYROID HORMONE-RELATED PEPTIDE. THE ACTIVITY OF THIS
 CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL
 CC CYCLASE AND ALSO A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER
 CC SYSTEM.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN
 CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

 CC EMBL; X78936; CAA55536.1; -;
 DR EMBL; L34611; AAA40011.1; -;
 DR EMBL; L34608; AAA40011.1; JOINED.
 DR EMBL; L34607; AAA40011.1; JOINED.
 DR EMBL; L34609; AAA40011.1; JOINED.
 DR EMBL; L34610; AAA40011.1; JOINED.
 DR GCRDB; GCR_1005; -;
 DR GCRDB; GCR_1614; -;
 DR MGD; MGI:97801; PTHR.
 DR INTERPRO; IPR000832; -;
 DR INTERPRO; IPR002170; -;
 DR PFAM; PF00002; 7tm_2; 1.
 DR PRINTS; PR00249; GPCRSECRETIN.
 DR PRINTS; PR00393; PTRHORMONER.
 DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
 DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 591
 FT DOMAIN 27 188
 FT TRANSMEM 189 212
 FT DOMAIN 214 219
 FT TRANSMEM 220 239
 FT DOMAIN 240 282
 FT TRANSMEM 283 306
 FT DOMAIN 307 320
 FT TRANSMEM 321 342
 FT DOMAIN 343 361
 FT TRANSMEM 362 382
 FT DOMAIN 383 409
 FT TRANSMEM 410 428
 FT DOMAIN 429 440
 FT TRANSMEM 441 463
 FT DOMAIN 464 591
 FT CARBOHYD 151 151
 FT CARBOHYD 161 161
 FT CARBOHYD 166 166
 FT CARBOHYD 176 176
 FT CONFLICT 27 28
 FT CONFLICT 464 465
 FT CONFLICT 500 501
 FT SEQUENCE 591 AA; 66313 MW; F7876F8D388BDDFD CRC64;

Query Match 51.6%; Score 79; DB 1; Length 591;
 Best Local Similarity 55.6%; Pred. No. 0.00063;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 09:03:42 ; Search time 58.45 Seconds
(without alignments)
13.665 Million cell updates/sec

Title: US-09-236-468a-2_COPY_56_80
Perfect score: 153
Sequence: 1 LQEGEGNCFPEWDGLICWRGTGK 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 87993

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	153	100.0	550	1 PTR2_HUMAN	P49190 homo sapien
2	145	94.8	546	1 PTH2_RAT	P70555 rattus norv
3	89	58.2	585	1 PTRR_DIDMA	P25107 didelphis m
4	79	51.6	591	1 PTRR_MOUSE	P41593 mus musculu
5	79	51.6	591	1 PTRR_RAT	P25961 rattus norv
6	76	49.7	585	1 PTRR_PIG	P50133 sus scrofa
7	76	49.7	593	1 PTRR_HUMAN	Q03431 homo sapien
8	64	41.8	423	1 GRFR_MOUSE	P32082 mus musculu
9	64	41.8	464	1 GRFR_RAT	Q02644 rattus norv
10	61	39.9	423	1 GRFR_HUMAN	Q02643 homo sapien
11	61	39.9	423	1 GRFR_PIG	P34999 sus scrofa
12	59.5	38.9	464	1 CGRR_RAT	Q63118 rattus norv
13	59	38.6	474	1 CALR_RABIT	P79222 oryctolagus
14	56.5	36.9	441	1 DIHR_ACHDO	Q16983 acheta dome
15	56.5	36.9	458	1 VIPR_PIG	Q28992 sus scrofa
16	56	36.6	478	1 CALR_CAVPO	O08893 cavia porce
17	56	36.6	516	1 CALR_RAT	P32214 rattus norv
18	55	35.9	515	1 CALR_MOUSE	Q60755 mus musculu
19	54	35.3	490	1 CALR_HUMAN	P30988 homo sapien
20	53.5	35.0	457	1 VIPR_HUMAN	P32241 homo sapien
21	53.5	35.0	461	1 CGRR_HUMAN	Q16602 homo sapien
22	53	34.6	445	1 SCRC_RABIT	O46502 oryctolagus
23	53	34.6	523	1 PACR_RAT	P32215 rattus norv
24	52	34.0	447	1 VIPR_CARAU	Q90308 carassius a
25	52	34.0	459	1 VIPR_RAT	P30083 rattus norv
26	52	34.0	496	1 PACR_MOUSE	P70205 mus musculu
27	51	33.3	213	1 ILLI_HUMAN	P15814 homo sapien
28	51	33.3	440	1 SCRC_HUMAN	P47872 homo sapien
29	51	33.3	498	1 CALR_PIG	P25117 sus scrofa
30	50	32.7	437	1 VIPR_MOUSE	Q41588 mus musculu
31	49.5	32.4	756	1 PRLI_HUMAN	P29674 homo sapien
32	49	32.0	513	1 PACR_BOVIN	Q29627 bos taurus
33	48	31.4	265	1 CXBA_BACVW	Q45790 bacillus th

RESULT	1
PTR2_HUMAN	
ID	PTR2_HUMAN STANDARD; PRT; 550 AA.
AC	P49190;
DT	01-FEB-1996 (Rel. 33, Created)
DI	01-FEB-1996 (Rel. 33, Last sequence update)
DT	01-OCT-2000 (Rel. 40, Last annotation update)
DE	PARATHYROID HORMONE RECEPTOR PRECURSOR (PTH2 RECEPTOR).
GN	PTH2.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN	[1]
RP	SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC	TISSUE=BRAIN;
RX	MEDLINE; 95318121.
RA	Usdin T.B., Gruber C., Bonner T.I.;
RT	"Identification and functional expression of a receptor selectively recognizing parathyroid hormone, the PTH2 receptor.";
RT	J. Biol. Chem. 270:15455-15458(1995).
RN	[2]
RP	SEQUENCE OF 26-40 AND 306-550 FROM N.A.
RX	MEDLINE; 97079671.
RA	Usdin T.B., Modi W., Bonner T.I.;
RT	"Assignment of the human PTH2 receptor gene (PTH2) to chromosome 2q33 by fluorescence in situ hybridization.";
RL	Genomics 37:140-141(1996)
CC	-!- FUNCTION: THIS IS A SPECIFIC RECEPTOR FOR PARATHYROID HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYL CYCLASE.
CC	-!- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN BRAIN AND PANCREAS.
CC	ALSO EXPRESSED IN THE TESTIS.
CC	-!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
DR	EMBL; U25128; AAC50157.1; -
DR	EMBL; U47124; AAA96796.1; -
DR	EMBL; U47129; AAC50767.1; -
DR	EMBL; U47125; AAC50767.1; JOINED.
DR	EMBL; U47126; AAC50767.1; JOINED.
DR	EMBL; U47127; AAC50767.1; JOINED.
DR	EMBL; U47128; AAC50767.1; JOINED.
DR	CCRD; GCR 2003; -
DR	MIM; 601469; -
DR	INTERPRO; IPR000832; -
DR	PFAM; PF00002; 7um_2; 1.
DR	PRINTS; PR00249; GPCRSECRETIN.
DR	PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR	PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.

THIS PAGE BLANK (USPTO)

Query Match 35.9%; Score 55; DB 2; Length 515;
Best Local Similarity 52.2%; Pred. No. 3.1;
Matches 12; Conservative 1; Mismatches 6; Indels 4; Gaps 2;

Qy 3 EGEG-NCFPEWDGLICW---PRG 21
||| | ||| :|| |
Db 66 EGGLYCNRTWDGWCDDTPAG 88

Search completed: November 8, 2000, 08:53:22
Job time: 353 sec

A:Molecule type: mRNA
A:Residues: 1-593 <LEV>
A:Cross-references: EMBL:U17418; NID:G596129; PIDN:AAA56774.1; PID:G596130
C:Genetics:
A:Introns: 25/3; 60/1; 105/1; 142/1; 181/3; 213/2; 278/3; 330/1; 350/2; 372/3; 404/2; 450/1
C:Superfamily: glucagon receptor
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 49.7%; Score 76; DB 2; Length 593;
Best Local Similarity 55.6%; Pred. No. 0.0038;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 8 CPPEWDGLICWPRGTGK 25
I I I I I I I I I I I I
Db 108 CLPEWDHILCWPLGAPGE 125

RESULT 7
S29753
growth hormone-releasing factor receptor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Nov-1999
C:Accession: S29753
R:Lin, C.; Lin, S.C.; Chang, C.P.; Rosenfeld, M.G.
Nature 360, 765-768, 1992
A:Title: Pit-1-dependent expression of the receptor for growth hormone releasing factor
A:Reference number: S29753; MUID:93101218
A:Accession: S29753
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-423 <LIN>
A:Cross-references: EMBL:L07379
C:Superfamily: glucagon receptor
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 41.8%; Score 64; DB 2; Length 423;
Best Local Similarity 48.0%; Pred. No. 0.14;
Matches 12; Conservative 1; Mismatches 6; Indels 6; Gaps 1;

QY 1 LOEGEG-----NCPPEWDGLICWP 19
I I I I I I I I I I I I
Db 42 LQAAGGTNTSLCCPGTWDGLLWCP 66

RESULT 8
S29754
growth hormone-releasing factor receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 05-Nov-1999
C:Accession: S29754; B45363
R:Lin, C.; Lin, S.C.; Chang, C.P.; Rosenfeld, M.G.
Nature 360, 765-768, 1992
A:Title: Pit-1-dependent expression of the receptor for growth hormone releasing factor
A:Reference number: S29753; MUID:93101218
A:Accession: S29754
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-464 <LIN>
A:Cross-references: EMBL:L07380
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, February 1993
R:Mayo, K.E.
Mol. Endocrinol. 6, 1734-1744, 1992
A:Title: Molecular cloning and expression of a pituitary-specific receptor for growth hormone releasing factor
A:Reference number: A45363; MUID:93078807
A:Accession: B45363
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-232, 'S', 234-304, 'C', 306-324, 366-464 <MAY>
A:Cross-references: GB:L01407; NID:G204314; PIDN:AAA41221.1; PID:G204315
A:Note: sequence extracted from NCBI backbone (NCBIP:118824)
C:Superfamily: glucagon receptor

Query Match 41.8%; Score 64; DB 2; Length 464;
Best Local Similarity 48.0%; Pred. No. 0.15;
Matches 12; Conservative 1; Mismatches 6; Indels 6; Gaps 1;

QY 1 LOEGEG-----NCPPEWDGLICWP 19
I I I I I I I I I I I I
Db 42 LQAAGGTNTSSMGCPGTWDGLLWCP 66

RESULT 9
A45363
somatoliberin receptor precursor - human
N:Alternate names: growth hormone-releasing hormone receptor (GHRH-R)
C:Species: Homo sapiens (man)
C:Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000
C:Accession: A45363; A45367
R:Mayo, K.E.
Mol. Endocrinol. 6, 1734-1744, 1992
A:Title: Molecular cloning and expression of a pituitary-specific receptor for growth hormone releasing factor
A:Reference number: A45363; MUID:93078807
A:Accession: A45363
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-423 <MAY>
A:Cross-references: GB:L01406; NID:G183172; PIDN:AAA35890.1; PID:G183173
A:Note: sequence extracted from NCBI backbone (NCBIP:118822)
R:Gaylinn, B.D.; Harrison, J.K.; Zysk, J.R.; Lyons, C.E.; Lynch, K.R.; Thorner, M.O.
Mol. Endocrinol. 7, 77-84, 1993
A:Title: Molecular cloning and expression of a human anterior pituitary receptor for growth hormone releasing factor
A:Reference number: A45367; MUID:93188867
A:Accession: A45367
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-177, 'A', 179-245, 'I', 247-423 <GAY>
A:Cross-references: PIDN:AAB25577.1; PID:G298370
A:Experimental source: acromegalic pituitary tumor
A:Note: sequence extracted from NCBI backbone (NCBIP:127815)
C:Genetics:
A:Gene: GDB:GHRHR
A:Cross-references: GDB:138465; OMIM:139191
A:Map position: 7p14-7p147p15-7p14
C:Superfamily: glucagon receptor
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-423/Product: somatoliberin receptor #status predicted <MAT>
F:50,216/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 39.9%; Score 61; DB 2; Length 423;
Best Local Similarity 50.0%; Pred. No. 0.37;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 8 CPPEWDGLICWPRGTGK 25
I I I I I I I I I I I I
Db 55 CPATWDGLLWPTAGSGE 72

RESULT 10
I46586
growth hormone-releasing hormone receptor - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 24-Nov-1999
C:Accession: I46586
R:Hsiung, H.; Smith, D.P.; Zhang, X.
NeuroReport 8, 1-10, 1997
A:Title: Structure and functional expression of a complementary DNA for porcine growth hormone releasing factor
A:Reference number: I46586; MUID:94020094
A:Accession: I46586
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-451 <HSI>

THIS PAGE BLANK (USPTO)

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 418 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-946-232-8

```

Query Match 39.9%; Score 61; DB 1; Length 418;
Best Local Similarity 50.0%; Pred. No. 0.45;
Matches 9; Conservative 2; Mismatches 7; Indels

Qy 8 CFPEWDGLICWPRGTVGK 25
+ + + + + + + + + +
Db 55 CPATWDGLLCWPTAGSGE 72

RESULT 15

US-08-073-799C-8
; Sequence 8, Application US/08073799C
; Patent No. 5583010

```

: GENERAL INFORMATION.
:
: APPLICANT: Zysk, John R.
: APPLICANT: Baumbach, William R.
: APPLICANT: Thorne, Michael O.
: APPLICANT: Gaylinn, Bruce D.
: APPLICANT: Lynch, Kevin R.
: APPLICANT: Harrison, Jeffrey K.
: TITLE OF INVENTION: ISOLATION,
: TITLE OF INVENTION: PROTEIN SPO
: TITLE OF INVENTION: AND CLONING
: TITLE OF INVENTION: HORMONE REC
: NUMBER OF SEQUENCES: 11
:

```

/ CORRESPONDENCE ADDRESS:
 /
 / ADDRESSEE: POPHAM, HAIK, SCHNOBRICH & KAUFMAN, LTD.
 /
 / STREET: Metropolitan Square Building, Suite 800, 1450 G St., N.W.
 /
 / CITY: Washington
 /
 / STATE: D.C.
 /
 / COUNTRY: U.S.A.
 /
 / 7 TD. 20005

```

;
;      DATE: 06-20-2003
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: PatentIn Release #1.0. version #1.25

```

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/073,799C
; FILING DATE: June 10, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;

```

APPLICATION NUMBER: US 07/902,826
FILING DATE: 23-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: O'Shaughnessy, Brian P.
REGISTRATION NUMBER: 32,747
REFERENCE/DOCKET NUMBER: 1574/81-
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-824-8000
TELEFAX: 202-824-8199
TELEX: 248516
INFORMATION SOURCE ID NO. 8

```

; INFORMATION FOR SEQ ID NO: 8
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 423 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-073-799C-8

```

Query Match 39.9%; Score 61; DB 1; Length 423;
Best Local Similarity 50.0%;
Matches 9; Conservative 2; Mismatches 7; Indels

QY	8	CFPEWDGLICWPRGTVGK	25
Db	55	CPATWDGLLCWPAGSGE	72

Search completed: November 8, 2000, 08:49:16
Job time: 109 sec

;; APPLICANT: Lyons Jr., Charles E.
;; TITLE OF INVENTION: MOLECULAR CLONING OF THE OVINE PITUITARY
;; GROWTH HORMONE RELEASING HORMONE RECEPTOR
;; NUMBER OF SEQUENCES: 18
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: POPHAM, HAIR, SCHNOBRICH & KAUFMAN, LTD.
;; STREET: Metropolitan Square Building, Suite 800, 1450
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20005
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/660,963
;; FILING DATE: 12-JUN-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: O'Shaughnessy, Brian P.
;; REGISTRATION NUMBER: 32,747
;; REFERENCE/DOCKET NUMBER: 18046.036
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-824-8000
;; TELEFAX: 202-824-8199
;; TELEX: 248516
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 457 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-660-963-10

Query Match 41.8%; Score 64; DB 2; Length 457;
Best Local Similarity 41.9%; Pred. No. 0.19;
Matches 13; Conservative 2; Mismatches 10; Indels 6; Gaps 1;

Qy 1 LQEGEG-----NCPPEWDGLICPRGTGK 25
11 11 11111111 1;
Db 92 LQAEGMPNSTLGCPRINDGLLCPMAGSGE 122

RESULT 13
US-08-660-963-11
; Sequence 11, Application us/08660963
; Patent No. 5852187
; GENERAL INFORMATION:
; APPLICANT: Thorne, Michael O.
; APPLICANT: Gaylinn, Bruce D.
; APPLICANT: Horikawa, Reiko
; APPLICANT: Lyons Jr., Charles E.
; TITLE OF INVENTION: MOLECULAR CLONING OF THE OVINE PITUITARY
; GROWTH HORMONE RELEASING HORMONE RECEPTOR
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: POPHAM, HAIR, SCHNOBRICH & KAUFMAN, LTD.
; STREET: Metropolitan Square Building, Suite 800, 1450
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/660,963
;; FILING DATE: 12-JUN-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: O'Shaughnessy, Brian P.
;; REGISTRATION NUMBER: 32,747
;; REFERENCE/DOCKET NUMBER: 18046.036
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-824-8000
;; TELEFAX: 202-824-8199
;; TELEX: 248516
;; INFORMATION FOR SEQ ID NO: 11:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 513 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-660-963-11

Query Match 41.8%; Score 64; DB 2; Length 513;
Best Local Similarity 41.9%; Pred. No. 0.21;
Matches 13; Conservative 2; Mismatches 10; Indels 6; Gaps 1;

Qy 1 LQEGEG-----NCPPEWDGLICPRGTGK 25
11 11 11111111 1;
Db 92 LQAEGMPNSTLGCPRINDGLLCPMAGSGE 122

RESULT 14
US-07-946-232-8
; Sequence 8, Application US/07946232
; Patent No. 5591641
; GENERAL INFORMATION:
; APPLICANT: Thorne, Michael O.
; APPLICANT: Gaylinn, Bruce D.
; APPLICANT: Lynch, Kevin R.
; APPLICANT: Harrison, Jeffrey K.
; APPLICANT: Zysk, John R.
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND
; PROTEIN SEQUENCING OF THE GROWTH HORMONE RELEASING
; HORMONE RECEPTOR AND CLONING OF A GENE ENCODING FOR THE
; GROWTH HORMONE RELEASING HORMONE RECEPTOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MASON, FENWICK & LAWRENCE
; STREET: 1225 I Street, N.W., Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,232
; FILING DATE: 19920915
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/902,826
; FILING DATE: 23-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Schein, Daniel B.
; REGISTRATION NUMBER: 33,551
; REFERENCE/DOCKET NUMBER: 1084/81-1317
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-289-1200
; TELEFAX: 202-289-6674
; TELEX: 248516
; INFORMATION FOR SEQ ID NO: 8:

```

; FILING DATE: 04-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/071003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-249A-19

Query Match 58.2%; Score 89; DB 2; Length 585;
Best Local Similarity 57.1%; Pred. No. 8.5e-05;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 5 EGNCPPEWDGLICWPRGTGK 25
: | | | | | : | | | | |
Db 102 DGECLPEWDNIVCWPGVPEK 122

RESULT 10
US-08-468-249A-20
; Sequence 20, Application US/08468249A
; Patent No. 5886148
; GENERAL INFORMATION:
; APPLICANT: Segre et al., Gino V.
; TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR AND DNA
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,249A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,475
; FILING DATE: 06-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,702
; FILING DATE: 04-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/071003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-249A-20

```

```

Query Match 51.6%; Score 79; DB 2; Length 591;
Best Local Similarity 55.6%; Pred. No. 0.0021;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 8 CPEWDGLICWPRGTGK 25
: | | | | | : | | | | |
Db 108 CLPEWDNIVCWPLGAPGE 125

RESULT 11
US-08-468-249A-21
; Sequence 21, Application US/08468249A
; Patent No. 5886148
; GENERAL INFORMATION:
; APPLICANT: Segre et al., Gino V.
; TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR AND DNA
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,249A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,475
; FILING DATE: 06-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,702
; FILING DATE: 04-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/071003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-249A-21

Query Match 49.7%; Score 76; DB 2; Length 593;
Best Local Similarity 55.6%; Pred. No. 0.0054;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 8 CPEWDGLICWPRGTGK 25
: | | | | | : | | | | |
Db 108 CLPEWDHILCWPLGAPGE 125

RESULT 12
US-08-660-963-10
; Sequence 10, Application US/08660963
; Patent No. 5852187
; GENERAL INFORMATION:
; APPLICANT: Thorne, Michael O.
; APPLICANT: Gaylinn, Bruce D.
; APPLICANT: Horikawa, Reiko

```

ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: 699 Prince Street
CITY: Alexandria
STATE: Virginia
COUNTRY: US
ZIP: 22313
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142,551B
FILING DATE: 25-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/077,296
FILING DATE: 14-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/898,219
FILING DATE: 12-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,677
FILING DATE: 22-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Swiss, Gerald F.
REGISTRATION NUMBER: 30,113
REFERENCE/DOCKET NUMBER: 000324-010
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 854-7400
TELEFAX: (415) 854-8275
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..585
OTHER INFORMATION: /note= "PTH receptor"
US-08-142-551B-125

Query Match 58.2%; Score 89; DB 2; Length 585;
Best Local Similarity 57.1%; Pred. No. 8.5e-05;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 5 EGNCFPEWDGLICWPRGTGK 25
: | | | | | : | | | | |
Db 102 DGFCLEPDNDIVCWPAGVPGK 122

RESULT 8
US-08-869-477-6
Sequence 6, Application US/08869477
Patent No. 5846747
GENERAL INFORMATION:
APPLICANT: Thorens, Bernard
TITLE OF INVENTION: Receptor for the Glucagon-Like-Peptide-1
TITLE OF INVENTION: (GLP-1)
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5846747o No. 5846747disk of No. 5846747th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/869,477
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/142,439
FILING DATE: 24-NOV-93
APPLICATION NUMBER: DK 398/92
FILING DATE: 25-MAR-92
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP93/00697
FILING DATE: 23-MAR-93
ATTORNEY/AGENT INFORMATION:
NAME: Harrington, James J.
REGISTRATION NUMBER: 38,711
REFERENCE/DOCKET NUMBER: 3756.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Didelphis virginiana
US-08-869-477-6

Query Match 58.2%; Score 89; DB 2; Length 585;
Best Local Similarity 57.1%; Pred. No. 8.5e-05;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 5 EGNCFPEWDGLICWPRGTGK 25
: | | | | | : | | | | |
Db 102 DGFCLEPDNDIVCWPAGVPGK 122

RESULT 9
US-08-468-249A-19
Sequence 19, Application US/08468249A
Patent No. 5886148
GENERAL INFORMATION:
APPLICANT: Segre et al., Gino V.
TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR AND DNA
TITLE OF INVENTION: ENCODING SAME
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,249A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,475
FILING DATE: 06-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/681,702

; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-011A-16

Query Match 58.2%; Score 89; DB 3; Length 60;
Best Local Similarity 57.1%; Pred. No. 7.2e-06;

Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 5 EGNCFPEWDGLICWPRGTGK 25

Db 1 DGFCLPEWDNVWCWPGVPGK 21

RESULT 5

US-08-468-249A-18
; Sequence 18, Application US/08468249A
; Patent No. 5886148

; GENERAL INFORMATION:
; APPLICANT: Segre et al., Gino V.

; TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR AND DNA

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/468,249A

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/864,475

; FILING DATE: 06-APR-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/681,702

; FILING DATE: 04-MAY-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Fraser, Janis K.

; REGISTRATION NUMBER: 34,819

; REFERENCE/DOCKET NUMBER: 00786/071003

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; INFORMATION FOR SEQ ID NO: 18:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 515 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-468-249A-18

Query Match 58.2%; Score 89; DB 2; Length 515;
Best Local Similarity 57.1%; Pred. No. 7.4e-05;

Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 5 EGNCFPEWDGLICWPRGTGK 25

Db 102 DGFCLPEWDNVWCWPGVPGK 122

RESULT 6

US-08-142-439A-6

; Sequence 6, Application US/08142439A

; Patent No. 5670360

; GENERAL INFORMATION:

; APPLICANT: Thorens, Bernard

; TITLE OF INVENTION: Receptor for the Glucagon-Like-Peptide-1

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESS: No. 56703600 No. 5670360disk of No. 5670360th America, Inc.

; STREET: 405 Lexington Avenue, Suite 6400

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10174-6201

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/142,439A

; FILING DATE: 24-NOV-93

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: DK 398/92

; FILING DATE: 25-MAR-92

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP93/00697

; FILING DATE: 23-MAR-93

; ATTORNEY/AGENT INFORMATION:

; NAME: Harrington, James J.

; REGISTRATION NUMBER: 38,711

; REFERENCE/DOCKET NUMBER: 3756.204-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212 867 0123

; TELEFAX: 212 867 0298

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 585 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Didelphis virginiana

; US-08-142-439A-6

Query Match 58.2%; Score 89; DB 1; Length 585;

Best Local Similarity 57.1%; Pred. No. 8.5e-05;

Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 5 EGNCFPEWDGLICWPRGTGK 25

Db 102 DGFCLPEWDNVWCWPGVPGK 122

RESULT 7

US-08-142-551B-125

; Sequence 125, Application US/08142551B

; Patent No. 5814603

; GENERAL INFORMATION:

; APPLICANT: Oldenburg, Kevin R.

; APPLICANT: Selick, Harold E.

; TITLE OF INVENTION: COMPOUNDS WITH PTH ACTIVITY AND

; NUMBER OF SEQUENCES: 132

; CORRESPONDENCE ADDRESS:

Db 56 LOEGEGNCFPEWDGLICWPRTGVK 80

RESULT 2

PCT-US95-07085-2
; Sequence 2, Application PC/TUS9507085
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone Receptor
; TITLE OF INVENTION: HLTGDG74
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07085
; FILING DATE: 05-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-393
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-07085-2

Query Match 100.0%; Score 153; DB 4; Length 541;
Best Local Similarity 100.0%; Pred. No. 1.le-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LOEGEGNCFPEWDGLICWPRTGVK 25
|||||
Db 56 LOEGEGNCFPEWDGLICWPRTGVK 80

RESULT 3

US-08-468-011A-15
; Sequence 15, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTGDG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland

; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-011A-15

Query Match 86.9%; Score 133; DB 3; Length 60;
Best Local Similarity 100.0%; Pred. No. 5.8e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 EGNCFPPEWDGLICWPRTGVK 25
|||||
Db 1 EGNCFPPEWDGLICWPRTGVK 21

RESULT 4

US-08-468-011A-16
; Sequence 16, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTGDG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:49:15 ; Search time 97.15 Seconds
(without alignments)
4.313 Million cell updates/sec

Title: US-09-236-468A-2_COPY_56_80

Perfect score: 153
Sequence: 1 LQEGEGNCFPEWDGLICWPRGTGK 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 164575

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	153	100.0	541	3	US-08-468-011A-2
2	153	100.0	541	4	PCT-US95-07085-2
3	133	86.9	60	3	US-08-468-011A-15
4	89	58.2	60	3	US-08-468-011A-16
5	89	58.2	515	2	US-08-468-249A-18
6	89	58.2	585	1	US-08-142-439A-6
7	89	58.2	585	2	US-08-142-551B-125
8	89	58.2	585	2	US-08-869-477-6
9	89	58.2	585	2	US-08-468-249A-19
10	79	51.6	591	2	US-08-468-249A-20
11	76	49.7	593	2	US-08-468-249A-21
12	64	41.8	457	2	US-08-660-963-10
13	64	41.8	513	2	US-08-660-963-11
14	61	39.9	418	1	US-07-946-232-8
15	61	39.9	423	1	US-08-073-799C-8
16	61	39.9	423	1	US-07-947-672-8
17	61	39.9	423	1	US-08-432-043-8
18	61	39.9	423	2	US-08-660-963-8
19	57	37.3	367	1	US-08-075-533-4
20	57	37.3	367	2	US-08-948-176-4
21	57	37.3	367	4	PCT-US91-09160-4
22	56.5	36.9	458	1	US-08-112-817C-2
23	54	35.3	180	1	US-08-453-742-25
24	54	35.3	180	1	US-08-454-464-25
25	54	35.3	180	1	US-08-453-222-25
26	54	35.3	180	1	US-08-452-802-25
27	54	35.3	474	1	US-08-453-742-2
28	54	35.3	474	1	US-08-454-464-2

29	54	35.3	474	1	US-08-453-222-2	Sequence 2, Appli
30	54	35.3	474	1	US-08-452-802-2	Sequence 2, Appli
31	54	35.3	477	1	US-08-453-742-27	Sequence 27, Appl
32	54	35.3	477	1	US-08-454-464-27	Sequence 27, Appl
33	54	35.3	477	1	US-08-453-222-27	Sequence 27, Appl
34	54	35.3	477	1	US-08-452-802-27	Sequence 27, Appl
35	53	34.6	448	2	US-08-811-897A-18	Sequence 18, Appl
36	53	34.6	448	2	US-08-855-213-18	Sequence 18, Appl
37	53	34.6	467	2	US-08-811-897A-19	Sequence 19, Appl
38	53	34.6	467	2	US-08-855-213-19	Sequence 19, Appl
39	53	34.6	476	2	US-08-811-897A-20	Sequence 20, Appl
40	53	34.6	476	2	US-08-855-213-20	Sequence 20, Appl
41	53	34.6	495	2	US-08-811-897A-21	Sequence 21, Appl
42	53	34.6	495	2	US-08-855-213-21	Sequence 21, Appl
43	52	34.0	1324	2	US-08-811-897A-56	Sequence 56, Appl
44	51	33.3	482	1	US-07-792-885A-1	Sequence 1, Appli
45	51	33.3	482	1	US-08-142-439A-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-08-468-011A-2
; Sequence 2, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carelia, Byrne, Bain, Gilfillan, Cecchi,
; STREET: Stewart & Olstein
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-011A-2

Query Match 100.0%; Score 153; DB 3; Length 541;
Best Local Similarity 100.0%; Pred. No. 1.1e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LQEGEGNCFPEWDGLICWPRGTGK 25
|||||

SQ Sequence 477 AA;

Query Match 43.5%; Score 66.5; DB 15; Length 477;
Best Local Similarity 50.0%; Pred. No. 0.14;
Matches 12; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

Qy 3 EGEKNCFFPEWDGLICW---PRGTV 23
||||| | ||| : || | | :

Db 65 egepcnrtwdgwmcwddtpgvm 88

RESULT 15

R62447
ID R62447 standard; protein; 515 AA.

AC R62447;

XX 21-JUN-1995 (first entry)

DE Rat calcitonin receptor C1b protein.

XX Calcitonin receptor C1b.

XX Rattus rattus.

Key Location/Qualifiers

FT Domain 1..144
FT /note= "extracellular portion"
FT Domain 212..256
FT /note= "extracellular portion"
FT Domain 315..333
FT /note= "extracellular portion"
FT Domain 400..410
FT /note= "extracellular portion"
FT Domain 145..168
FT /note= "transmembrane portion"
FT Domain 185..211
FT /note= "transmembrane portion"
FT Domain 257..280
FT /note= "transmembrane portion"
FT Domain 293..314
FT /note= "transmembrane portion"
FT Domain 334..356
FT /note= "transmembrane portion"
FT Domain 380..399
FT /note= "transmembrane portion"
FT Domain 411..430
FT /note= "transmembrane portion"
FT Domain 169..184
FT /note= "intracellular portion"
FT Domain 281..292
FT /note= "intracellular portion"
FT Domain 357..379
FT /note= "intracellular portion"
FT Domain 431..515
FT /note= "intracellular portion"

PN WO9421665-A.

XX 29-SEP-1994.

XX 24-MAR-1994; 94WO-US03248.

XX 24-MAR-1993; 93US-0036784.

XX (AMYL-) AMYLIN PHARM INC.

XX Albrandt K, Beaumont K;

XX WPI; 1994-316927/39.

XX New calcitonin receptor proteins - used to identify agents for

PT treating e.g. obesity, anorexia, pain, diabetes mellitus or
PT insulin resistance

PS Disclosure; Fig. 1; 59pp; English.

XX The receptor protein can be used for determining the presence or
amount of or separating C1b receptor binding compounds in a
sample. They can also be used for producing antibodies. The
receptor protein is especially used for identifying calcitonin,
amylin or CGRP agonists or antagonists for treating conditions such
as obesity, anorexia, pain, diabetes mellitus impaired glucose
tolerance or insulin resistance.

XX Sequence 515 AA;

Query Match 43.5%; Score 66.5; DB 15; Length 515;
Best Local Similarity 50.0%; Pred. No. 0.15;
Matches 12; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

Qy 3 EGEKNCFFPEWDGLICW---PRGTV 23

Db 65 egepcnrtwdgwmcwddtpgvm 88

Search completed: November 8, 2000, 08:51:40
Job time: 252 sec

Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 8 CFPWDGLICWPRGTGK 25
| | | | : | | | | |
Db 108 clpewdhilcwpplgapge 125

RESULT 13

R27707
ID R27707 standard; Protein; 614 AA.

XX

AC R27707;

XX 16-MAR-1993 (first entry)

XX Human kidney PTH/PTHrP receptor.

XX Parathyroid hormone; related protein; calcium; antagonist;
KW antibodies; hypercalcaemia.

XX Iiomo sapiens.

XX WO9217602-A.

XX 15-OCT-1992.

XX 06-APR-1992; 92WO-US02821.

XX 05-APR-1991; 91US-0681702.

XX 06-APR-1992; 92US-0864475.

XX (GEO) GEN HOSPITAL CORP OFFICE TECHNOLOGY.

XX Abou-samra A, Juppner H, Kronenberg HM, Potts JT, Schipani E;
PI Segre GV;

XX WPI; 1992-366271/44.

XX N-PSDB; Q29607.

XX New DNA encoding parathyroid hormone receptor, DNA and antibodies
PT - for (differential) diagnosis of hypercalcaemia, and diagnosis
PT and treatment of tumours
XX Claim 22; Fig 6; 91pp; English.

XX The protein sequence was deduced from the cDNA sequence obt'd by
CC screening a human kidney oligo dt-primed cDNA library in lambda
CC gt10 and a genomic library of human placental DNA in EMBl3 with a
CC probe comprising most of the coding sequence of rat bone parathyroid
CC hormone/parathyroid hormone related protein (PTH/PTHrP) receptor
CC protein. The clone encodes a protein which may be used in a
CC therapeutic compsn. to inhibit activation of PTH or PTHrP and thus
CC reduce the level of calcium in the blood. Cpd's. capable of competing
CC with PTH or PTHrP for binding can be identified using the protein prod.
CC and DNAs homologous to PTH DNA can be identified using fragments of the
CC clone as probes. The sequence may be used for the prodn. of antibodies
CC useful for the treatment, classification, prognosis and/or treatment of
CC disorders related to the interaction between a cell receptor and a
CC ligand such as in hypercalcaemia. See also R27704-16.

XX Sequence 614 AA;

Query Match 46.4%; Score 71; DB 13; Length 614;
Best Local Similarity 64.3%; Pred. No. 0.043;

Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 8 CFPWDGLICWPRG 21

| | | | : | | | | |

Db 108 clpewdhilcwpplg 121

RESULT 14

R62448

ID R62448 standard; protein; 477 AA.

XX

AC R62448;

XX 21-JUN-1995 (first entry)

XX Rat calcitonin receptor Cla protein.

XX Calcitonin receptor Cla.

XX Rattus rattus.

XX Key

XX Location/Qualifiers

XX 1..144

XX /note= "extracellular portion"

XX 212..219

XX /note= "extracellular portion"

XX 278..296

XX /note= "extracellular portion"

XX 363..373

XX /note= "extracellular portion"

XX 145..168

XX /note= "transmembrane portion"

XX 185..211

XX /note= "transmembrane portion"

XX 220..243

XX /note= "transmembrane portion"

XX 256..277

XX /note= "transmembrane portion"

XX 297..319

XX /note= "transmembrane portion"

XX 343..362

XX /note= "transmembrane portion"

XX 374..393

XX /note= "transmembrane portion"

XX 169..184

XX /note= "intracellular portion"

XX 244..255

XX /note= "intracellular portion"

XX 320..342

XX /note= "intracellular portion"

XX 394..477

XX /note= "intracellular portion"

XX WO9421665-A.

XX 29-SEP-1994.

XX 24-MAR-1994; 94WO-US03248.

XX 24-MAR-1993; 93US-0036784.

XX (AMYL-) AMYLIN PHARM INC.

XX Albrandt K, Beaumont K;

XX WPI; 1994-316927/39.

XX New calcitonin receptor proteins - used to identify agents for

XX treating e.g. obesity, anorexia, pain, diabetes mellitus or

XX insulin resistance

XX Disclosure; Fig. 1; 59pp; English.

XX The receptor protein can be used for determining the presence or

XX amount of or separating Clb receptor binding compounds in a

XX sample. They can also be used for producing antibodies. The

XX receptor protein is especially used for identifying calcitonin,

XX amylin or CGRP agonists or antagonists for treating conditions such

XX as obesity, anorexia, pain, diabetes mellitus impaired glucose

XX tolerance or insulin resistance.

XX

PR 06-JUN-1995; 95US-0471494.
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 XX
 PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;
 XX
 DR WPI: 1999-034124/03.
 DR N-PSDB: V08390.
 XX
 PT Antibody to parathyroid hormone receptor - for diagnostic or
 PT therapeutic use
 XX
 XX
 PS Claim 6; Fig 3; 63pp; English.
 XX
 CC This sequence represents the rat parathyroid hormone (PTH) receptor
 CC R15b, which is targeted by the antibody of the invention. The antibody
 CC of the invention is immunoreactive with naturally occurring human, rat or
 CC opossum PTH receptor. The antibody is useful for treating disorders
 CC characterised by overstimulation of PTH receptors by their ligand and for
 CC the diagnosis of PTH-related hypercalcaemia.
 XX
 SQ Sequence 591 AA;
 Query Match 51.6%; Score 79; DB 20; Length 591;
 Best Local Similarity 55.6%; Pred. No. 0.0033;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 8 CPPEWDGLICWPRGVGK 25
 Db 108 clpewdnivcplgapge 125
 RESULT 11
 R92278
 ID R92278 standard; Protein; 593 AA.
 XX
 AC R92278;
 XX
 DT 18-MAY-1996 (first entry)
 XX
 DE Human kidney PTH/PTHrP receptor.
 XX
 KW Parathyroid hormone; receptor; parathormone; PTH;
 KW parathyroid hormone-related protein; PTHrP; calcium; homeostasis;
 KW hypercalcaemia; hypocalcaemia; cancer.
 XX
 OS Homo sapiens.
 XX
 PN US5494806-A.
 XX
 PD 27-FEB-1996.
 XX
 PE 05-APR-1991; 91US-0681702.
 XX
 PR 06-APR-1992; 92US-0864475.
 PR 05-APR-1991; 91US-0681702.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;
 XX
 DR WPI: 1996-139028/14.
 DR N-PSDB; R15948.
 XX
 PT DNA encoding vertebrate parathyroid hormone receptor - useful for
 PT diagnosis and treatment of e.g. hypercalcaemia, hypo-calcaemia,
 PT cancer etc.
 XX
 PS Claim 1; Fig 6A-6G; 64pp; English.
 XX

CC A human parathyroid hormone/parathyroid hormone-related protein
 CC (PTH/PTHrP) receptor (R92278) is encoded by cDNA clone HK-1
 CC (R15948) isolated from a human kidney cDNA library. The receptor
 CC induces an increase in intracellular cAMP and intracellular free
 CC calcium when challenged by PTH or PTHrP. Recombinant receptor can
 CC be produced in vector/host cell systems and used in the treatment,
 CC diagnosis or prognosis of e.g. hypercalcaemia and hypocalcaemia, to
 CC screen for (ant)agonists and to raise antibodies. Host cells
 CC expressing the receptor can be used for diagnostic measurement of
 CC PTH serum levels.
 XX
 XX Sequence 593 AA;
 SQ
 Query Match 49.7%; Score 76; DB 17; Length 593;
 Best Local Similarity 55.6%; Pred. No. 0.0085;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 8 CPPEWDGLICWPRGVGK 25
 Db 108 clpewdhilcwpplgapge 125
 RESULT 12
 W73317
 ID W73317 standard; Protein; 593 AA.
 XX
 AC W73317;
 XX
 DT 08-FEB-1999 (first entry)
 XX
 DE Human Parathyroid hormone receptor.
 XX
 KW Parathyroid hormone receptor; PTH receptor; antibody; therapy;
 KW PTH-related hypercalcaemia; human.
 XX
 OS Homo sapiens.
 XX
 PN US5840853-A.
 XX
 PD 24-NOV-1998.
 XX
 PE 06-JUN-1995; 95US-0471494.
 XX
 PR 06-APR-1992; 92US-0864475.
 PR 05-APR-1991; 91US-0681702.
 PR 06-JUN-1995; 95US-0471494.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;
 XX
 DR WPI: 1999-034124/03.
 DR N-PSDB; V08391.
 XX
 PT Antibody to parathyroid hormone receptor - for diagnostic or
 PT therapeutic use
 XX
 XX Claim 7; Fig 6; 63pp; English.
 XX
 CC This sequence represents the human parathyroid hormone (PTH) receptor
 CC which is targeted by the antibody of the invention. The antibody
 CC of the invention is immunoreactive with naturally occurring human, rat or
 CC opossum PTH receptor. The antibody is useful for treating disorders
 CC characterised by overstimulation of PTH receptors by their ligand and for
 CC the diagnosis of PTH-related hypercalcaemia.
 XX
 XX Sequence 593 AA;
 SQ
 Query Match 49.7%; Score 76; DB 20; Length 593;
 Best Local Similarity 55.6%; Pred. No. 0.0085;

CC The rat bone parathyroid hormone/parathyroid hormone related
 CC protein (PTHrP/PTHrP) receptor protein sequence was deduced from
 CC clone R15B obtd. by screening a rat osteosarcoma (ROS) cell cDNA library
 CC to isolate those expressing functionally intact PTH/PTHrP receptor
 CC proteins, performed according to Gearing et al., (EMBO J. 8: 3676,
 CC 1989), by identifying colonies capable of binding a suitable radio-
 CC labelled ligand. The protein may be used in a therapeutic compsn. to
 CC inhibit activation of PTH or PTHrP and thus reduce the level of calcium
 CC in the blood. Cdds. capable of competing with PTH or PTHrP for binding
 CC can be identified using the protein and DNAs homologous to PTH DNA can
 CC be identified using fragments of the clone as probes. The sequence
 CC may be used for the prodn. of antibodies useful for the treatment,
 CC classification, prognosis and/or treatment of disorders related to
 CC the interaction between a cell receptor and a ligand such as in
 CC hypercalcaemia. See also R27704-16.

XX Sequence 591 AA;

Query Match 51.6%; Score 79; DB 13; Length 591;
 Best Local Similarity 55.6%; Pred. No. 0.0033;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 8 CPPEWDGLICWPRGTGVGK 25
 | |||| : || | | :
 Db 108 clpewdnivcwpplgagpe 125

RESULT 9
 R92277 ID R92277 standard; Protein; 591 AA.

XX R92277;
 XX
 XX
 DT 18-MAY-1996 (first entry)

XX Rat bone PTH/PTHrP receptor.

XX Parathyroid hormone; receptor; parathormone; PTH;
 KW parathyroid hormone-related protein; PTHrP; calcium; homeostasis;
 KW hypercalcaemia; hypocalcaemia; cancer.

OS Rattus sp.

Key	Location/Qualifiers
FT Region	1..192
FT	/label- Extracellular_region
FT Region	193..211
FT	/label- Transmembrane_region
FT Region	212..221
FT	/label- Intracellular_region
FT Region	222..240
FT	/label- Transmembrane_region
FT Region	241..299
FT	/label- Extracellular_region
FT Region	300..316
FT	/label- Transmembrane_region
FT Region	317..325
FT	/label- Intracellular_region
FT Region	326..342
FT	/label- Transmembrane_region
FT Region	343..364
FT	/label- Extracellular_region
FT Region	365..383
FT	/label- Transmembrane_region
FT Region	384..408
FT	/label- Intracellular_region
FT Region	409..428
FT	/label- Transmembrane_region
FT Region	429..444
FT	/label- Intracellular_region
FT Region	445..463
FT	/label- Transmembrane_region

FT Region 464..591
 /label- Intracellular_region
 FT
 XX US5494806-A.
 PN
 XX 27-FEB-1996.
 PD
 XX
 XX 05-APR-1991; 91US-0681702.
 PF
 XX 06-APR-1992; 92US-0864475.
 PR 05-APR-1991; 91US-0681702.
 XX
 XX (GEO) GEN HOSPITAL CORP.
 PA
 XX Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;
 XX
 DR WPI; 1996-139028/14.
 DR N-PSDB; T15947.
 XX
 XX DNA encoding vertebrate parathyroid hormone receptor - useful for
 PT diagnosis and treatment of e.g. hypercalcaemia, hypo-calcaemia,
 PT cancer etc.

XX Claim 1; Fig 3A-3E; 64pp; English.

XX A rat parathyroid hormone/parathyroid hormone-related protein
 CC (PTH/PTHrP) receptor (R92277) is encoded by cDNA clone R15B
 CC (T15947) isolated from a rat osteosarcoma ROS 17/2.8 cDNA library.
 CC The receptor a G-protein linked receptor having 7 transmembrane
 CC domains. It induces an increase in intracellular cAMP and calcium
 CC upon challenge with PTH or PTHrP. Recombinant receptor can be
 CC produced in vector/host cell systems and used in the treatment,
 CC diagnosis or prognosis of e.g. hypercalcaemia and hypocalcaemia,
 CC to screen for (ant)agonists and to raise antibodies. Host cells
 CC expressing the receptor are used for diagnostic measurement of PTH
 CC serum levels.

XX Sequence 591 AA;

Query Match 51.6%; Score 79; DB 17; Length 591;
 Best Local Similarity 55.6%; Pred. No. 0.0033;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 8 CPPEWDGLICWPRGTGVGK 25
 | |||| : || | | :
 Db 108 clpewdnivcwpplgagpe 125

RESULT 10
 W73316 ID W73316 standard; Protein; 591 AA.

XX W73316;
 AC
 XX
 XX 08-FEB-1999 (first entry)
 DT
 XX
 DE Parathyroid hormone receptor R15B.
 XX
 KW Parathyroid hormone receptor; PTH receptor; antibody; therapy;
 KW PTH-related hypercalcaemia; rat.

XX Rattus sp.

XX US5840853-A.
 PN
 XX 24-NOV-1998.
 PD
 XX 06-JUN-1995; 95US-0471494.
 PF
 XX 06-APR-1992; 92US-0864475.
 PR 05-APR-1991; 91US-0681702.

KW hypercalcaemia; hypocalcaemia; cancer; opossum.
 OS Didelphis virginiana.
 PN US5494806-A.
 XX 27-FEB-1996.
 PD
 PF 05-APR-1991; 91US-0681702.
 XX
 XX 06-APR-1992; 92US-0864475.
 PR 05-APR-1991; 91US-0681702.
 XX (GEHO) GEN HOSPITAL CORP.
 PA
 XX Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;
 XX WPI; 1996-139028/14.
 DR N-PSDB; T15946.
 XX
 PT DNA encoding vertebrate parathyroid hormone receptor - useful for
 PT diagnosis and treatment of e.g. hypercalcaemia, hypo-calcaemia,
 PT cancer etc.
 XX
 FS Claim 1; Fig 2A-2E; 64pp; English.
 XX
 CC Parathyroid hormone/parathyroid hormone-related protein (PTH/PTHrP)
 CC receptors (R92275 and R92776) are encoded by cDNA clones OK-H (T15945)
 CC and OK-O (T15946), respectively, isolated from an opossum kidney (OK)
 CC cell cDNA library. OK-H and OK-O may be the products of 2 separate
 CC genes or of a laboratory artifact. The receptor induces an increase
 CC in intracellular cAMP and calcium when challenged with PTH or PTHrP.
 CC Recombinant receptors can be produced in vector/host cell systems and
 CC used in the treatment, diagnosis or prognosis of e.g. hypercalcaemia
 CC and hypocalcaemia, to screen for (ant)agonists and to raise antibodies.
 CC Host cells expressing the receptor can be used for diagnostic
 CC measurement of PTH serum levels.
 XX
 SQ Sequence 585 AA;
 Query Match 58.2%; Score 89; DB 17; Length 585;
 Best Local Similarity 57.1%; Pred. No. 0.00013;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 Qy 5 EGNCFPEWDGLICWPRGTVGK 25
 :| | | | | :| | | | |
 Db 102 dgfcfpewdnivcwpagvpgk 122
 RESULT 7
 W73315
 ID W73315 standard; Protein; 585 AA.
 XX
 AC W73315;
 XX
 XX 08-FEB-1999 (first entry)
 DT
 DE Parathyroid hormone receptor OK-O.
 XX
 KW Parathyroid hormone receptor; PTH receptor; antibody; therapy;
 KW PTH-related hypercalcaemia; opossum.
 XX
 OS Didelphis virginiana.
 XX
 PN US5840853-A.
 XX
 XX 24-NOV-1998.
 PD
 XX 06-JUN-1995; 95US-0471494.
 PF
 XX 06-APR-1992; 92US-0864475.
 PR

PR 05-APR-1991; 91US-0681702.
 PR 06-JUN-1995; 95US-0471494.
 XX (GEHO) GEN HOSPITAL CORP.
 PA
 XX Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;
 XX WPI; 1999-034124/03.
 DR N-PSDB; V08389.
 XX
 PT Antibody to parathyroid hormone receptor - for diagnostic or
 PT therapeutic use
 XX
 FS Claim 6; Fig 2; 63pp; English.
 XX
 CC This sequence represents the opossum parathyroid hormone (PTH) receptor
 CC OK-O, which is targeted by the antibody of the invention. The antibody
 CC of the invention is immunoreactive with naturally occurring human, rat or
 CC opossum PTH receptor. The antibody is useful for treating disorders
 CC characterised by overstimulation of PTH receptors by their ligand and for
 CC the diagnosis of PTH-related hypercalcaemia.
 XX
 SQ Sequence 585 AA;
 Query Match 58.2%; Score 89; DB 20; Length 585;
 Best Local Similarity 57.1%; Pred. No. 0.00013;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 Qy 5 EGNCFPEWDGLICWPRGTVGK 25
 :| | | | | :| | | | |
 Db 102 dgfcfpewdnivcwpagvpgk 122
 RESULT 8
 R27706
 ID R27706 standard; Protein; 591 AA.
 XX
 AC R27706;
 XX
 XX 16-MAR-1993 (first entry)
 DT
 XX Rat bone PTH/PTHrP receptor clone R15B prod.
 DE
 XX Parathyroid hormone; related protein; calcium; antagonist;
 KW antibodies; hypercalcaemia.
 KW
 XX Rattus rattus.
 OS
 XX W09217602-A.
 PN
 XX 15-OCT-1992.
 PD
 XX 06-APR-1992; 92WO-US02821.
 PF
 XX 05-APR-1991; 91US-0681702.
 PR 06-APR-1992; 92US-0864475.
 XX
 XX (GEHO) GEN HOSPITAL CORP OFFICE TECHNOLOGY.
 PA
 XX Abou-samra A, Juppner H, Kronenberg HM, Potts JT, Schipani E;
 PI Segre GV;
 XX WPI; 1992-366271/44.
 DR N-PSDB; Q29606.
 DR
 XX New DNA encoding parathyroid hormone receptor, DNA and antibodies
 PT - for (differential) diagnosis of hypercalcaemia, and diagnosis
 PT and treatment of tumours
 XX
 XX Claim 20; Fig 3; 91pp; English.
 PS
 XX

Query Match 58.2%; Score 89; DB 17; Length 515;
 Best Local Similarity 57.1%; Pred. No. 0.00012;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 5 EGNCFPEWDGLICWPRGTGVGK 25
 :||| |||| :||| | ||
 Db 102 dgfcipewdnivcwpagvpvk 122

RESULT 4
 W73314
 ID W73314 standard; Protein; 515 AA.

XX AC W73314;

XX DT 08-FEB-1999 (first entry)

XX DE Parathyroid hormone receptor OK-H.

XX KW Parathyroid hormone receptor; PTH receptor; antibody; therapy;
 KW PTH-related hypercalcaemia; opossum.

XX OS Didelphis virginiana.

XX PN US5840853-A.

XX PD 24-NOV-1998.

XX PF 06-JUN-1995; 95US-0471494.

XX PR 06-APR-1992; 92US-0864475.

XX PR 05-APR-1991; 91US-0681702.

XX PR 06-JUN-1995; 95US-0471494.

XX PA (GEO) GEN HOSPITAL CORP.

XX PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;

XX PI Schipani E, Segre GV;

XX DR WPI; 1999-034124/03.

XX DR N-PSDB; V08386.

XX PT Antibody to parathyroid hormone receptor - for diagnostic or

XX PT therapeutic use

XX PS Claim 6; Fig 1; 63pp; English.

XX CC This sequence represents the opossum parathyroid hormone (PTH) receptor
 CC OK-H, which is targeted by the antibody of the invention. The antibody
 CC of the invention is immunoreactive with naturally occurring human, rat or
 CC opossum PTH receptor. The antibody is useful for treating disorders
 CC characterised by overstimulation of PTH receptors by their ligand and for
 CC the diagnosis of PTH-related hypercalcaemia.

XX SQ Sequence 515 AA;

Query Match 58.2%; Score 89; DB 20; Length 515;
 Best Local Similarity 57.1%; Pred. No. 0.00012;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 5 EGNCFPEWDGLICWPRGTGVGK 25
 :||| |||| :||| | ||
 Db 102 dgfcipewdnivcwpagvpvk 122

RESULT 5
 R27705
 ID R27705 standard; Protein; 585 AA.

XX AC R27705;

XX DT 16-MAR-1993 (first entry)

XX DE Opossum kidney PTH/PTHrP receptor prod. from clone OK-O.
 XX KW Parathyroid hormone; related protein; calcium; antagonist;
 XX KW antibodies; hypercalcaemia.

XX OS Didelphis virginiana.

XX PN W09217602-A.

XX PD 15-OCT-1992.

XX PF 06-APR-1992; 92WO-US02821.

XX PR 05-APR-1991; 91US-0681702.

XX PR 06-APR-1992; 92US-0864475.

XX PA (GEO) GEN HOSPITAL CORP OFFICE TECHNOLOGY.

XX PI Abou-samra A, Juppner H, Kronenberg HM, Potts JT, Schipani E;

XX PI Segre GV;

XX DR WPI; 1992-366271/44.

XX DR N-PSDB; Q29605.

XX PT New DNA encoding parathyroid hormone receptor, DNA and antibodies
 PT - for (differential) diagnosis of hypercalcaemia, and diagnosis
 PT and treatment of tumours

XX PS Disclosure; Fig 2; 91pp; English.

XX CC The parathyroid hormone/parathyroid hormone related protein (PTH/PTHrP)
 CC receptor protein sequence was deduced from the DNA sequence of the
 CC clone OK-O, isolated from opossum kidney (OK) cells. The clone OK-O
 CC is identical to the OK-H clone except at the C-terminal tail as OK-O
 CC encodes a 585 amino acid protein, OK-H encodes one of 515 amino acids.
 CC The difference is attributed to a single nucleotide deleted in the OK-H
 CC sequence causing a frame shift and an earlier stop codon. It is not
 CC known whether OK-O and OK-H represent prods. of two separate genes or
 CC are a laboratory artifact. The protein may be used in a therapeutic
 CC compsn. to inhibit activation of PTH or PTHrP and thus reduce the
 CC level of calcium in the blood. Cpd. capable of competing with PTH
 CC or PTHrP for binding can be identified using the protein prod. and
 CC DNAs homologous to PTH DNA can be identified using fragments of the
 CC clone as probes. The sequence may be used for the prodn. of antibodies
 CC useful for the treatment, classification, prognosis and/or treatment of
 CC disorders related to the interaction between a cell receptor and a
 CC ligand such as in hypercalcaemia. See also R27704-16.

XX SQ Sequence 585 AA;

Query Match 58.2%; Score 89; DB 13; Length 585;
 Best Local Similarity 57.1%; Pred. No. 0.00013;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 5 EGNCFPEWDGLICWPRGTGVGK 25
 :||| |||| :||| | ||
 Db 102 dgfcipewdnivcwpagvpvk 122

RESULT 6
 R92276
 ID R92276 standard; Protein; 585 AA.

XX AC R92276;

XX DT 18-MAY-1996 (first entry)

XX DE Opossum kidney PTH/PTHrP receptor.

XX KW Parathyroid hormone; receptor; parathormone; PTH;

XX KW parathyroid hormone-related protein; PTHrP; calcium; homeostasis;

PT hyper-calcaemia, hypo- or hyper-phosphatemia, kidney stones, etc
 PS Claim 9; Fig 1A-E; 62pp; English.
 CC A novel 7-transmembrane receptor (W12695) has been identified as a
 CC human G-protein parathyroid hormone (PTH) receptor, designated as a
 CC HLTDC74. It shows 48.2% homology to the human PTH receptor. Its
 CC amino acid sequence was deduced from a cDNA clone (T59619) isolated
 CC from a human T cell lymphoma tissue cDNA library. Recombinant
 CC HLTDC74 can be produced in transfected host cells and used to
 CC screen for (ant)agonist cpds. Agonists may be used to prevent or
 CC and chronic tetany by stimulating an increase in serum calcium
 CC levels. Antagonists can be used to inhibit the receptor e.g. for
 CC the treatment of osteoporosis, hypercalcaemia, hypoparathyroidism,
 CC hypophosphataemia, kidney stone, nephrolisis.
 XX Sequence 541 AA;
 SQ

Query Match 100.0%; Score 153; DB 18; Length 541;
 Best Local Similarity 100.0%; Pred. No. 1.7e-13;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQEGEGNCFPEWDGLICWPRGTGK 25
 |||||||||
 Db 56 lqegegncfpewdglicwprgtvgk 80

RESULT 2
 R27704
 ID R27704 standard; Protein; 515 AA.
 XX
 AC R27704;
 XX
 DT 16-MAR-1993 (first entry)
 XX
 DE Opossum kidney PTH/PTHrP receptor prep. from clone OK-H.
 XX
 KW Parathyroid hormone; related protein; calcium; antagonist;
 KW antibodies; hypercalcaemia.
 OS Didelphis virginiana.
 XX
 PN W09217602-A.
 XX
 PD 15-OCT-1992.
 XX
 PF 06-APR-1992; 92WO-US02821.
 XX
 PR 05-APR-1991; 91US-0681702.
 PR 06-APR-1992; 92US-0864475.
 XX
 XX (GEO) GEN HOSPITAL CORP OFFICE TECHNOLOGY.
 PA
 PI Abou-samra A, Juppner H, Kronenberg HM, Potts JT, Schipani E;
 PI Segre GV;
 XX
 DR WPI; 1992-366271/44.
 DR N-PSDB; Q29604.
 XX
 XX New DNA encoding parathyroid hormone receptor, DNA and antibodies
 PT - for (differential) diagnosis of hypercalcaemia, and diagnosis
 PT and treatment of tumours
 XX
 PS Disclosure; Fig 1; 91pp; English.
 XX
 CC The parathyroid hormone/parathyroid hormone related protein (PTH/PTHrP)
 CC receptor protein sequence was deduced from the DNA sequence of the
 CC clone OK-H, isolated from opossum kidney (OK) cells. The protein
 CC may be used in a therapeutic compsn. to inhibit activation of PTH or
 CC PTHrP and thus reduce the level of calcium in the blood. Cpds.
 CC capable of competing with PTH or PTHrP for binding can be identified

CC using the protein prod. and DNAs homologous to PTH DNA can be
 CC identified using fragments of the clone as probes. The protein
 CC may be used for the prodn. of antibodies useful for the treatment,
 CC classification, prognosis and/or treatment of disorders related to
 CC the interaction between a cell receptor and a ligand such as in
 CC hypercalcaemia. See also R27703-16.
 XX
 SQ Sequence 515 AA;
 Query Match 58.2%; Score 89; DB 13; Length 515;
 Best Local Similarity 57.1%; Pred. No. 0.00012;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 5 EGNCFPEWDGLICWPRGTGK 25
 :|||:|:|:|:|:|:|
 Db 102 dgfcipewdnivcwpagvpgk 122

RESULT 3
 R92275
 ID R92275 standard; Protein; 515 AA.
 XX
 AC R92275;
 XX
 DT 18-MAY-1996 (first entry)
 XX
 DE Opossum kidney PTH/PTHrP receptor.
 XX
 KW Parathyroid hormone; receptor; parathormone; PTH;
 KW parathyroid hormone-related protein; PTHrP; calcium; homeostasis;
 KW hypercalcaemia; hypocalcaemia; cancer; opossum.
 XX
 OS Didelphis virginiana.
 XX
 PN US5494806-A.
 XX
 PD 27-FEB-1996.
 XX
 PF 05-APR-1991; 91US-0681702.
 XX
 PR 06-APR-1992; 92US-0864475.
 PR 05-APR-1991; 91US-0681702.
 XX
 XX (GEO) GEN HOSPITAL CORP.
 PA
 PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;
 XX
 DR WPI; 1996-139028/14.
 DR N-PSDB; T15945.
 XX
 PT DNA encoding vertebrate parathyroid hormone receptor - useful for
 PT diagnosis and treatment of e.g. hypercalcaemia, hypo-calcaemia,
 PT cancer etc.
 XX
 PS Claim 1; Fig 1A-1E; 64pp; English.
 XX
 CC Parathyroid hormone/parathyroid hormone-related protein (PTH/PTHrP)
 CC receptors (R92275 and R92776) are encoded by cDNA clones OK-H (T15945)
 CC and OK-O (T15946), respectively, isolated from an opossum kidney (OK)
 CC cell cDNA library. OK-H and OK-O may be the products of 2 separate
 CC genes or of a laboratory artifact. The receptor induces an increase
 CC in intracellular cAMP and calcium when challenged with PTH or PTHrP.
 CC Recombinant receptors can be produced in vector/host cell systems and
 CC used in the treatment, diagnosis or prognosis of e.g. hypercalcaemia
 CC and hypocalcaemia, to screen for (ant)agonists and to raise antibodies.
 CC Host cells expressing the receptor can be used for diagnostic
 CC measurement of PTH serum levels.
 XX
 SQ Sequence 515 AA;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:51:39 ; Search time 138.73 Seconds
(without alignments)
6.162 Million cell updates/sec

Title: US-09-236-468A-2_COPY_56_80

Perfect score: 133

Sequence: 1 LQEGEGNCFFPDGLICWPRGTGK 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_36.*

- 1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	153	100.0	541	18 W12695	G-protein parathyr
2	89	58.2	515	13 R27704	Opessum kidney PTH
3	89	58.2	515	17 R92275	Opessum kidney PTH
4	89	58.2	515	20 W73314	Parathyroid hormon
5	89	58.2	585	13 R27705	Opessum kidney PTH
6	89	58.2	585	17 R92276	Opessum kidney PTH
7	89	58.2	585	20 W73315	Parathyroid hormon
8	79	51.6	591	13 R27706	Rat bone PTH/PTHrP
9	79	51.6	591	17 R92277	Rat bone PTH/PTHrP
10	79	51.6	591	20 W73316	Parathyroid hormon
11	76	49.7	593	17 R92278	Human kidney PTH/P
12	76	49.7	593	20 W73317	Human Parathyroid

13	71	46.4	514	13 R27707	Human kidney PTH/P
14	66.5	43.5	477	15 R62448	Rat calcitonin rec
15	66.5	43.5	515	15 R62447	Rat calcitonin rec
16	64	41.8	423	15 R51844	Rat somatoliberin
17	64	41.8	457	20 W67743	Ovine growth hormo
18	61	39.9	418	15 R51071	Human Growth Hormo
19	61	39.9	423	15 R51843	Human somatoliberi
20	61	39.9	423	15 R53801	G3 pGRF receptor.
21	61	39.9	423	15 R51072	Human Growth Hormo
22	61	39.9	423	18 W24033	Growth hormone-rel
23	61	39.9	423	18 W10654	Human growth hormo
24	61	39.9	423	18 W10488	Human acromegalic
25	61	39.9	423	20 W67744	Human clone HAP7.3
26	61	39.9	429	15 R53800	pGRF receptor. Su
27	61	39.9	451	15 R53799	pGRF receptor. Su
28	59	38.6	423	18 W10487	Porcine growth hor
29	57	37.3	367	20 Y28633	Soybean seed acyl-
30	56.5	36.9	458	16 R72506	Porcine vasoactive
31	54	35.3	180	15 R51703	cDNA coding human
32	54	35.3	474	15 R51702	Human calcitonin r
33	54	35.3	477	15 R51704	Human placental ca
34	54	35.3	490	14 R37424	Human CTR. Homo s
35	53.5	35.0	461	17 R92220	Calcitonin gene-re
36	53	34.6	448	15 R58666	Rat PACAP receptor
37	53	34.6	467	15 R58656	Rat PACAP receptor
38	53	34.6	476	15 R58657	Rat PACAP receptor
39	53	34.6	495	15 R58658	Rat PACAP receptor
40	53	34.6	1381	18 W31947	Rat CASPR/pl90, a
41	53	34.6	1384	18 W31946	Human CASPR/pl90,
42	52	34.0	459	14 R42848	VIP receptor prote
43	51.5	33.7	293	21 Y75339	Neisseria meningit
44	50	32.7	437	19 W80309	Rat PACAP/VIP R-2
45	50	32.7	437	20 W92973	Rat PACAP/VIP R-2

ALIGNMENTS

RESULT	ID	W12695	standard; Protein; 541 AA.
XX	AC	W12695;	
XX	DT	31-MAY-1997 (first entry)	
XX	DE	G-protein parathyroid hormone receptor HLTG74.	

XX G-protein parathyroid hormone receptor; HLTG74; parathormone; PTH:
KW calcium; signal transduction; agonist; antagonist; hypocalcaemia;
KW hyperphosphataemia; hypoparathyroidism; chronic tetany;
KW osteoporosis; hypercalcaemia; hypoparathyroidism; hypophosphataemia;
KW kidney stone; nephroliasis; therapy; diagnosis.

OS Homo sapiens.

PN WO9639433-A1.

PD 12-DEC-1996.

PF 05-JUN-1995; 95WO-US07085.

PR 05-JUN-1995; 95WO-US07085.

PA (HUMA-) HUMAN GENOME SCI INC.

XX Li Y, Rosen CA, Ruben SM, Soppet DR;

PI WPI: 1997-043068/04.

DR N-PSDB; T59619.

XX Human G-protein parathyroid hormone receptor, HLTG74 - used to
PT identify (ant)agonists, used in the treatment of hypo- or


```

ID O45049 PRELIMINARY; PRT; 494 AA.
AC O45049;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-1998 (TREMBlrel. 06, Last annotation update)
DE INOSITOL TRISPHOSPHATE 3-KINASE FORM 1.
GN LFE-2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA Claudiin T.R.; Demodena J.A.; Sternberg P.W.;
RL Cell 0:0-0(1998).
DR EMBL; AF045611; AAC38960.1; -.
SQ SEQUENCE 494 AA; 56784 MW; 8F1EBEC9F0359F89 CRC64;

Query Match 29.0%; Score 51; DB 5; Length 494;
Best Local Similarity 55.6%; Pred. No. 24;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 8 EDGGRQDDILMEKPSRP 25
|:|:| | | | | | | | | |
Db 99 EESGHHPSDIQEKPNRP 116

RESULT 13
Q9QZJ5 PRELIMINARY; PRT; 595 AA.
AC Q9QZJ5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE ESTROGEN RECEPTOR ALPHA.
GN ER-ALPHA.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
RN [1]
RP SEQUENCE FROM N.A.
RA Bhat H.K.; Vadgama J.V.;
RT "Hamster estrogen receptor cDNA: cloning and mRNA expression.";
RL J. Steroid Biochem. Mol. Biol. 0:0-0(1999).
DR EMBL; AF181077; AAD53956.1; -.
DR INTERPRO; IPR000324; -.
DR INTERPRO; IPR000536; -.
DR INTERPRO; IPR001292; -.
DR INTERPRO; IPR001628; -.
DR INTERPRO; IPR001723; -.
DR PFAM; PF00104; hormone_rec; 1.
DR PFAM; PF00105; zf-C4; 1.
DR PRINTS; PR00047; STROIDFINGER.
DR PRINTS; PR00350; VITAMINDR.
DR PRINTS; PR00398; STRDHORMONER.
DR PRINTS; PR00543; OESTROGENR.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor.
SQ SEQUENCE 595 AA; 66833 MW; E4D803B4FFDB257E CRC64;

Query Match 29.0%; Score 51; DB 11; Length 595;
Best Local Similarity 41.9%; Pred. No. 30;
Matches 13; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 2 STRSNKEDSGRQDDILMEKPSRPSPDPT 32
| | | | | | | | | | | | | |
Db 152 SNSNDRQSGRRRLSSSSSEKSGMAMESVKET 182

RESULT 14
Q9UEG4 PRELIMINARY; PRT; 927 AA.
AC Q9UEG4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE KIAA0326 PROTEIN (FRAGMENT).
GN KIAA0326.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue=Brain;
RA Nagase T.; Ishikawa K.; Nakajima D.; Ohira M.; Seki N.; Miyajima N.;
RA Tanaka A.; Kotani H.; Nomura N.; Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
DR EMBL; AB002324; BAA20784.1; -.
DR INTERPRO; IPR000822; -.
DR PFAM; PF00096; zf-C2H2; 19.
DR PRINTS; PR00048; ZINC_FINGER.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 19.
FT NON_TER 1
SQ SEQUENCE 927 AA; 102756 MW; 176C78A9C407F8C0 CRC64;

```

```

Query Match 28.7%; Score 50.5; DB 4; Length 927;
Best Local Similarity 38.7%; Pred. No. 55;
Matches 12; Conservative 7; Mismatches 9; Indels 3; Gaps 1;

QY 6 NKEDSGROR---DDILMEKPSRPSPDPT 33
|:|:| | | | | | | | | | | |
Db 88 NEESPROESSGEEIIMGDPQSPESKOSTE 118

RESULT 15
Q9VFF92 PRELIMINARY; PRT; 931 AA.
AC Q9VFF92;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE CG4913 PROTEIN.
GN CG4913.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RX MEDLINE; 20196006.
RA Adams M.D.; Celniker S.E.; Holt R.A.; Evans C.A.; Gocayne J.D.;
RA Amanatides P.G.; Scherer S.E.; Li P.W.; Hoskins R.A.; Galle R.F.;
RA George R.A.; Lewis S.E.; Richards S.; Ashburner M.; Henderson S.N.;
RA Sutton G.G.; Wortman J.R.; Yandell M.D.; Zhang Q.; Chen L.X.;
RA Brandon R.C.; Rogers Y.-H.C.; Blazew R.G.; Champe M.; Pfeiffer B.D.;
RA Wan K.H.; Doyle C.; Baxter E.G.; Helt G.; Nelson C.R.; Miklos G.L.G.;
RA Abril J.F.; Agbayani A.; An H.-J.; Andrews-Pfannkoch C.; Baldwin D.;
RA Ballew R.M.; Basu A.; Baxendale J.; Bayraktaroglu L.; Beasley E.M.;
RA Beson K.Y.; Benos P.V.; Berman B.P.; Blandari D.; Bolshakov S.;
RA Borkova D.; Botchan M.R.; Bouck J.; Brokstein P.; Brotter P.;
RA Burtis K.C.; Busam D.A.; Butler H.; Cadieu E.; Center A.; Chandra I.;
RA Cherry J.M.; Cawley S.; Dahlke C.; Davenport L.B.; Davies P.;
RA de Pablos B.; Delcher A.; Deng Z.; Mays A.D.; Dew I.; Dietz S.M.;
RA Dodson K.; Doup L.E.; Downes M.; Dugan-Rocha S.; Dunkov B.C.; Dunn P.;
RA Durbin K.J.; Evangelista C.C.; Ferraz C.; Ferreira S.; Fleischmann W.;
RA Foster C.; Gabrielian A.E.; Garg N.S.; Gelbart W.M.; Glasser K.;
RA Glodek A.; Gong F.; Gorrell J.H.; Gu Z.; Guan P.; Harris M.;
RA Harris N.L.; Harvey D.; Helman T.J.; Hernandez J.R.; Houck J.;
RA Hostin D.; Houston K.A.; Howland T.J.; Wei M.-H.; Ibegwam C.;

```


Query Match 29.5%; Score 52; DB 10; Length 834;
 Best Local Similarity 47.8%; Pred. No. 30;
 Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 5 SNKEDSGRQDDILMEKPSRPM 27
 II IIII: I I I :: I I
 DB 635 SNAEDSGMRDRDGLDSMDREPE 657

RESULT 9
 Q90ZG6 PRELIMINARY; PRT; 197 AA.
 AC Q90ZG6;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
 DE ESTROGEN ALPHA RECEPTOR (FRAGMENT).
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jones J.E., Carpenter C.D., Lubbers L.S., Petersen S.L., Wade G.N.;
 RT "Return of lordosis after food deprivation and refeeding in Syrian
 RL hamsters";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF184276; AAF02721.1; -;
 DR INTERPRO; IPR001292; -;
 DR INTERPRO; IPR001628; -;
 DR PFAM; PF00105; zf-C4; 1.
 DR PRINTS; PR00047; STROIDFINGER.
 DR PRINTS; PR00543; OESTROGENR.
 KW Receptor.
 FT NON_TER 1
 FT NON_TER 197 197
 SQ SEQUENCE 197 AA; 22435 MW; 65EE471F94D4A5F6 CRC64;

Query Match 29.0%; Score 51; DB 11; Length 197;
 Best Local Similarity 41.9%; Pred. No. 9.4;
 Matches 13; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 2 STRSNKEDSGRQDDILMEKPSRPMSPDT 32
 I I: IIII: I I I I I I I I
 DB 55 SNSDRRSGRRLSSSEKSGMAMESVKET 85

RESULT 10
 O53176 PRELIMINARY; PRT; 419 AA.
 ID O53176;
 AC O53176;
 DT 01-JUN-1998 (Tremblrel. 06, Created)
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
 DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
 DE HYPOTHETICAL 44.4 KDA PROTEIN.
 GN RV2449C OR MTV008.05C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE; 98295987.
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmler K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jageis K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,

RA Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RL complete genome sequence.";
 RL Nature 393:537-544(1998).
 DR EMBL; AL021246; CAA16026.1; -;
 DR TUBERCULIST; RV2449C; -;
 KW Hypothetical protein.
 SQ SEQUENCE 419 AA; 44378 MW; 83DA2104BC7DC630 CRC64;

Query Match 29.0%; Score 51; DB 2; Length 419;
 Best Local Similarity 48.1%; Pred. No. 21;
 Matches 13; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

QY 1 LSTRSNKEDSGRQDDILMEKPSRPM 27
 IIII I I I I I I I I I I
 DB 192 LSTASNDPDRQLSDPYMLSPDRGAE 218

RESULT 11
 P91166 PRELIMINARY; PRT; 486 AA.
 ID P91166;
 AC P91166;
 DT 01-MAY-1997 (Tremblrel. 03, Created)
 DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
 DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
 DE SIMILAR TO INOSITOL KINASES.
 GN C46H11.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE; 94150718.
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Faveilo A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RL elegans.";
 RN Nature 368:32-38(1994).
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Miller N., Bradshaw H., Wamsley P.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U88314; AAB42358.1; -;
 SQ SEQUENCE 486 AA; 56215 MW; 895336D893105F4F CRC64;

Query Match 29.0%; Score 51; DB 5; Length 486;
 Best Local Similarity 55.6%; Pred. No. 24;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 8 EDSGRQDDILMEKPSRP 25
 I: I I I I I I I I I I I I I I
 DB 84 EESGHHPSDIQNEKPNRP 101

RESULT 12
 O45049

aps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:56:22 ; Search time 152.43 Seconds
(without alignments)
20.827 Million cell updates/sec

Title: US-09-236-468A-2_COPY_508_541
Perfect score: 176
Sequence: 1 LSTRSNKEDSGRRDDILMEKPSRPMESNPDTREG 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues
Total number of hits satisfying chosen parameters: 297973

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:.*
1: sp_archaea:.*
2: sp_bacteria:.*
3: sp_fungi:.*
4: sp_human:.*
5: sp_invertebrate:.*
6: sp_mammal:.*
7: sp_mhc:.*
8: sp_organelle:.*
9: sp_phage:.*
10: sp_plant:.*
11: sp_rodent:.*
12: sp_virus:.*
13: sp_vertebrate:.*
14: sp_unclassified:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	57	32.4	575	13 Q9PWB7	Q9pwb7 brachydanio
2	56	31.8	568	12 O57285	O57285 sendai viru
3	54	30.7	575	4 Q9ULK5	Q9ulk5 homo sapien
4	54	30.7	1390	5 Q17602	Q17602 caenorhabdi
5	53	30.1	611	10 Q41042	Q41042 pisum sativ
6	53	30.1	1114	5 Q9VI80	Q9vi80 drosophila
7	53	30.1	1463	5 O44384	O44384 drosophila
8	52	29.5	834	10 Q9SVJ8	Q9svj8 arabidopsis
9	51	29.0	197	11 Q9QZG6	Q9qzg6 mesoicetetu
10	51	29.0	419	2 O53176	O53176 mycobacteri
11	51	29.0	486	5 P91166	P91166 caenorhabdi
12	51	29.0	494	5 O45049	O45049 caenorhabdi
13	51	29.0	595	11 Q9QZJ5	Q9qzj5 mesoicetetu
14	50.5	28.7	927	4 Q9UEG4	Q9ueg4 homo sapien
15	50.5	28.7	931	5 Q9VF92	Q9vf92 drosophila
16	50	28.4	288	5 Q9W0P0	Q9w0p0 drosophila
17	50	28.4	490	5 Q22852	Q22852 caenorhabdi
18	50	28.4	510	2 Q06818	Q06818 enterococcu
19	50	28.4	520	1 O26312	O26312 methanobact

20	50	28.4	1214	5	Q20129	Q20129 caenorhabdi
21	49.5	28.1	685	5	O76850	O76850 calliphora
22	49.5	28.1	918	10	O82206	O82206 arabidopsis
23	49.5	28.1	1121	10	O9ZPE1	O9zpe1 arabidopsis
24	49.5	28.1	2327	12	O9QCE4	O9qce4 foot and mo
25	49.5	28.1	2327	12	O9QCE3	O9qce3 foot and mo
26	49.5	28.1	2327	12	O9QCE2	O9qce2 foot and mo
27	49	27.8	95	6	O46505	O46505 equus cabal
28	49	27.8	177	5	O9VJ87	O9vj87 drosophila
29	49	27.8	395	1	O73956	O73956 haloferax s
30	49	27.8	416	5	O45514	O45514 caenorhabdi
31	49	27.8	718	2	O9Z5C8	O9z5c8 staphylococ
32	49	27.8	854	2	O9S3P8	O9s3p8 streptococc
33	49	27.8	862	2	O9RP22	O9rp22 streptococc
34	49	27.8	865	2	O9S4J9	O9s4j9 streptococc
35	49	27.8	872	2	O9S4J3	O9s4j3 streptococc
36	49	27.8	873	2	O9S4K0	O9s4k0 streptococc
37	49	27.8	881	2	O9RFJ4	O9rfj4 streptococc
38	49	27.8	923	2	O9S3T2	O9s3t2 streptococc
39	49	27.8	1025	2	O4507	O4507 streptococc
40	49	27.8	1025	2	P72532	P72532 streptococc
41	49	27.8	1029	2	O9X2V2	O9x2v2 streptococc
42	49	27.8	1225	5	Q17656	Q17656 caenorhabdi
43	48.5	27.6	862	3	O13414	O13414 aspergillus
44	48.5	27.6	2768	5	Q9VC00	Q9vc00 drosophila
45	48	27.3	250	5	O96207	O96207 plasmodium

ALIGNMENTS

RESULT 1	
Q9PWB7	
ID Q9PWB7	PRELIMINARY; PRT; 575 AA.
AC Q9PWB7	
DT 01-MAY-2000 (TREMBLrel. 13, Created)	
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)	
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)	
DE PARATHYROID HORMONE TYPE-2 RECEPTOR PRECURSOR.	
OS Brachydanio rerio (zebrafish) (zebra danio).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;	
OC Cypriniformes; Cyprinidae; Cyprinidae; Rasbora; Danio.	
RN [1]	
RP SEQUENCE FROM N.A.	
RC TISSUE=KIDNEY;	
RX MEDLINE; 99367425.	
RA Rubin D.A., Hellman P., Zon L.I., Lobb C.J., Bergwitz C., Juppner H.;	
RT "A G protein-coupled receptor from zebrafish is activated by human	
RT parathyroid hormone and not by human or teleost parathyroid hormone-	
RT related peptide. Implications for the evolutionary conservation of	
RT calcium-regulating peptide hormones."	
RL J. Biol. Chem. 274:23035-23042(1999).	
DR EMBL; AF132082; AAD51908.1; -	
DR INTERPRO; IPR000832; -	
DR INTERPRO; IPR001879; -	
DR PFAM; PF00002; 7tm_2; 1	
DR PRINTS; PR00249; GPCRSECRETIN.	
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.	
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.	
KW Receptor.	
SQ SEQUENCE 575 AA; 64244 MW; 888F1C4DD3A14DC CRC64;	

Query Match 32.4%; Score 57; DB 13; Length 575;
Best Local Similarity 46.4%; Pred. No. 4;
Matches 13; Conservative 5; Mismatches 8; Indels 2; Gaps 1;
QY 8 EDSGRQRDDILMEK--PSRPMESNPDTRE 33
|||:|||||:|:|
DB 545 EDSAKQVDDILKESLPTRPSSGLEDD 572

THIS PAGE BLANK (USPTO)

Qy 2 STRSNKEDSGQRDDI-----LMEK-----PSRPMESNP 30
:| ||| | ||| |:: | | |
Db 65 TTTTPGKEDEGELRDDIASLQGLLDRHVQFGRKLPLKTPYANP 106

Search completed: November 8, 2000, 09:04:00
Job time: 870 sec

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D63707; BAA09838.1; -.
DR MGB; MGI:1194494; HDGF.
DR INTERPRO; IPR000313; -.
DR PFAM; PF00855; PWMP; 1.
KW Growth factor; Heparin-binding.
FT DOMAIN 213 220 POLY-GLU.
SQ SEQUENCE 237 AA; 26278 MW; 60CC6BD74DA47320 CRC64;

Query Match 27.0%; Score 47.5; DB 1; Length 237;
Best Local Similarity 36.7%; Pred. No. 18;
Matches 11; Conservative 4; Mismatches 12; Indels 3; Gaps 1;

QY 7 KEDSG---RQRDDILMEKPSRPMESNPDT 33
| | | | | : | | | | | | | | | |
Db 148 KNEKGTLLKRRAGDVLDSPKRKESGDHEE 177

RESULT 14
SSK1_YEAST
ID SSK1_YEAST STANDARD; PRT; 712 AA.
AC Q07084; Q07909;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE OSMOLARITY TWO-COMPONENT SYSTEM PROTEIN SSK1.
GN SSK1 OR YLR006C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomycetes.
RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS.
RX MEDLINE; 94239498.
RA Maeda T., Wurgler-Murphy S.M., Saito H.;
RT "A two-component system that regulates an osmosensing MAP kinase
RT cascade in yeast."
RL Nature 369:242-245(1994).
RN [2]

CC SEQUENCE FROM N.A.
CC Vandenbol M., Portetelle D., Hilger F.;
CC Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: FORMS PART OF A TWO-COMPONENT REGULATORY SYSTEM
CC SSK1/SSK1 ACTIVATED BY CHANGES IN THE OSMOLARITY OF THE
CC EXTRACELLULAR ENVIRONMENT. THIS SYSTEM CONTROLS THE SSK2/SSK22->
CC PBS2->HOG1 PATHWAY. THE UNPHOSPHORYLATED FORM OF SSK1 ACTIVATES
CC SSK2 AND SSK22, TWO MAPKRS THAT FURTHER STIMULATE THE PBS2-HOG1
CC MAPK CASCADE. IN LOW-OSMOLARITY MEDIA, IS INHIBITED THROUGH THE
CC PHOSPHORYLATION BY SLN1.
CC -1- SUBUNIT: SEEMS TO INTERACT WITH SSK2 AND SSK22.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC -1- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; L26523; AAA35100.1; -.
DR EMBL; Z73178; CAA97528.1; -.
DR SGD; S0003996; SSK1.
DR INTERPRO; IPR001789; -.
KW Sensory transduction; Phosphorylation.
FT DOMAIN 501 646 RECEIVER DOMAIN (POTENTIAL).

Query Match 27.0%; Score 47.5; DB 1; Length 1082;
Best Local Similarity 33.3%; Pred. No. 1e+02;
Matches 14; Conservative 5; Mismatches 10; Indels 13; Gaps 2;

FT MOD_RES 554 554 PHOSPHORYLATION (PROBABLE).
FT MUTAGEN 554 D->N: ACTIVATES.
FT CONFLICT 181 P -> S (IN REF. 1).
SQ SEQUENCE 712 AA; 78529 MW; 33B2DBB4FCF2528A CRC64;

Query Match 27.0%; Score 47.5; DB 1; Length 712;
Best Local Similarity 36.0%; Pred. No. 64;
Matches 9; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

QY 7 KEDSGRQRDDILMEKPSRPM-ESNP 30
| | | | | : | | | | | | | | | |
Db 659 KQGESRWTDSDVIVKSPQPIAPSNP 683

RESULT 15
SP23_YEAST
ID SP23_YEAST STANDARD; PRT; 1082 AA.
AC P35210;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE SPT23 PROTEIN.
GN SPT23 OR YKL020C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomycetes.
RN [1]
RP SEQUENCE FROM N.A.
RA Rieger M.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 136-752 FROM N.A.
RX MEDLINE; 94262317.
RA Burkett T.J., Garfinkel D.J.;
RT "Molecular characterization of the SPT23 gene: a dosage-dependent
RT suppressor of ty-induced promoter mutations from Saccharomycetes
RT cerevisiae."
RL Yeast 10:81-92(1994).
CC -1- FUNCTION: DOSAGE-DEPENDENT SUPPRESSOR OF TY-INDUCED PROMOTER
CC MUTATIONS. MAY EXERT ITS SUPPRESSION EFFECT THROUGH PROTEIN-
CC PROTEIN INTERACTIONS SINCE DOES NOT PRESENT ANY OF THE MOTIFS
CC GENERALLY FOUND IN TRANSCRIPTIONAL ACTIVATORS OR DNA BINDING
CC PROTEINS.
CC -1- SIMILARITY: TO YEAST YIR033W.
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 738
CC ONWARD AND IS SHORTER (752 AA) DUE TO A FRAMESHIFT.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; Z28020; CAA81855.1; -.
DR EMBL; L24760; AAA20575.1; ALT_FRAME.
DR PIR; S37837; S37837.
DR SGD; S0001503; SPT23.
DR INTERPRO; IPR002110; -.
DR INTERPRO; IPR002909; -.
DR PFAM; PF01833; TIG; 1.
DR PFAM; PF00023; ank; 2.
DR PROSITE; PS50088; ANK_REPEAT; 2.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
FT CONFLICT 715 715 H -> P (IN REF. 2).
SQ SEQUENCE 1082 AA; 121337 MW; 620C688ECC0ECBD7 CRC64;

Query Match 27.0%; Score 47.5; DB 1; Length 1082;
Best Local Similarity 33.3%; Pred. No. 1e+02;
Matches 14; Conservative 5; Mismatches 10; Indels 13; Gaps 2;


```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01130; CAA24361.1; -.
DR EMBL; X00429; CAA25127.1; -.
DR PIR; A03908; GNNY2F.
DR HSSP; P13899; LTWE.
DR INTERPRO; IPR000605; -.
DR INTERPRO; IPR001205; -.
DR INTERPRO; IPR001676; -.
DR PFAM; PF00680; RNA_dep_RNA_pol; 1.
DR PFAM; PF00910; RNA_helicase; 1.
DR PFAM; PF00073; rhv; 3.
KW Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;
KW Transferase; Hydrolase; Thiol protease; Nonstructural protein;
KW Myristate.
FT CHAIN 1 201 NONSTRUCTURAL PROTEIN P20A.
FT CHAIN 202 286 COAT PROTEIN VP4.
FT CHAIN 287 504 COAT PROTEIN VP2.
FT CHAIN 505 725 COAT PROTEIN VP3.
FT CHAIN 726 937 COAT PROTEIN VP1.
FT CHAIN 938 1578 CORE PROTEIN P52.
FT CHAIN 1579 1601 GENOME-LINKED PROTEIN VP1.
FT CHAIN 1602 1625 GENOME-LINKED PROTEIN VP2.
FT CHAIN 1626 1649 GENOME-LINKED PROTEIN VP3.
FT CHAIN 1650 1863 PROTEASE P20B.
FT CHAIN 1864 2333 RNA-DIRECTED RNA POLYMERASE P56A.
FT LIPID 202 202 MYRISTATE.
FT CONFLICT 396 396 S -> C (IN REF. 2).
FT CONFLICT 632 632 P -> L (IN REF. 2).
SQ SEQUENCE 2333 AA; 259645 MW; 4FC667DCC521BC60 CRC64;

Query Match 28.1%; Score 49.5; DB 1; Length 2333;
Best Local Similarity 47.6%; Pred. No. 1.4e+02;
Matches 10; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

Qy 14 RDDILMEKPSRPMESNPDEG 34
| | | | | | | | | |
Db 1560 RDDVSSE-PAQPAEDQPAEG 1579

RESULT 10
HDGF_HUMAN STANDARD; PRT; 240 AA.
AC P51858;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HEPATOMA-DERIVED GROWTH FACTOR (HDGF).
GN HDGF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 4-24.
RC TISSUE=HEPATOMA;
RX MEDLINE; 95014294.
RA Nakamura H., Izumoto Y., Kambe H., Kuroda T., Mori T., Kawamura K.,
RA Yanamoto H., Kishimoto T.;
RT "Molecular cloning of complementary DNA for a novel human hepatoma-
RT derived growth factor. Its homology with high mobility group-1
RT protein.";
RL J. Biol. Chem. 269:25143-25149(1994).
CC -!- FUNCTION: HEPARIN-BINDING PROTEIN, WITH MITOGENIC ACTIVITY FOR
CC FIBROBLASTS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- TISSUE SPECIFICITY: UBQUITOUS.
CC -!- SIMILARITY: BELONGS TO THE HMGI/HMG2 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z73002; CAA97244.1; -.
DR EMBL; Z73003; CAA97245.1; -.

```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D16431; BAA03903.1; -.
DR MIM; 300043; -.
DR INTERPRO; IPR000313; -.
DR PFAM; PF00855; PWMP; 1.
KW Growth factor; Heparin-binding.
FT DOMAIN 213 223 POLY-GLU.
SQ SEQUENCE 240 AA; 26788 MW; DD60D9203BDD4B34 CRC64;

Query Match 27.8%; Score 49; DB 1; Length 240;
Best Local Similarity 38.7%; Pred. No. 11;
Matches 12; Conservative 6; Mismatches 9; Indels 4; Gaps 2;

Qy 7 KEDSG--RQDDILMEKPSRPMES-NPDYE 33
| | | | | | | | | |
Db 148 KNEGALKRRAGDLDLSPKRPKEAENPEGE 178

RESULT 11
CCHI_YEAST STANDARD; PRT; 2039 AA.
AC P50077;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PROBABLE CALCIUM-CHANNEL PROTEIN.
GN CCHI OR YGR217W OR G8501.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomycetes.
RN [1]
RP SEQUENCE OF 1-1360 FROM N.A.
RC STRAIN=S288C;
RX MEDLINE; 97435481.
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT chromosome VII.";
RL Yeast 13:1077-1090(1997).
RN [2]
RP SEQUENCE OF 1184-2039 FROM N.A.
RC STRAIN=S288C;
RX MEDLINE; 96267763.
RA van der Aart Q.J.M., Kleine K., Steensma H.Y.;
RT "Sequence analysis of the 43 Kb CRM1-YLM9-PET54-DIE2-SM11-PHO81-YHB4-
RT PFK1 region from the right arm of Saccharomyces cerevisiae chromosome
RT VII.";
RL Yeast 12:385-390(1996).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE; 98001560.
RA Paidhungat M., Garrett S.;
RT "A homolog of mammalian, voltage-gated calcium channels mediates
RT yeast pheromone-stimulated Ca2+ uptake and exacerbates the cdc1(Ts)
RT growth defect.";
RL Mol. Cell. Biol. 17:6339-6347(1997).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -!- SIMILARITY: SOME, TO MAMMALIAN SODIUM AND CALCIUM CHANNELS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z73002; CAA97244.1; -.
DR EMBL; Z73003; CAA97245.1; -.

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D49714; BAAL19916.1; -
DR MENDEL; 16305; Oryza; 1229; 16305.
DR INTERPRO; IPR000965; -
DR INTERPRO; IPR001048; -
DR INTERPRO; IPR001057; -
DR INTERPRO; IPR002086; -
DR PFAM; PF00696; aakinas; 1.
DR PFAM; PF00171; alidcdh; 1.
DR PRINTS; PR00474; GLU5KINASE.
DR PROSITE; PS00902; GLUTAMATE_5_KINASE; 1.
DR PROSITE; PS01223; PROA; 1.
KW Proline biosynthesis; Multifunctional enzyme; Oxidoreductase; NADP;
KW Transferase; Kinase.
FT DOMAIN 1 296 GLUTAMATE 5-KINASE.
FT DOMAIN 297 716 GAMMA-GLUTAMYL PHOSPHATE REDUCTASE.
FT SEQUENCE 716 AA; 77697 MW; C6D928D7C8D4A06E CRC64;

Query Match 29.0%; Score 51; DB 1; Length 716;
Best Local Similarity 33.3%; Pred. No. 22;
Matches 15; Conservative 6; Mismatches 10; Indels 14; Gaps 2;

QY 1 LSTRNKEDSGRQ-----RDILMEKPSRPW-----ENPD 31
DB 373 IRTLANMEDPINQILKKTAEVADDLVLEKTSCLPLGVILLIVFESRPD 417

RESULT 8
VGLL_HCMVA
ID VGLL_HCMVA STANDARD; PRT; 487 AA.
AC P16778; P87887;
DT 01-AUG-1990 (Rel. 15, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IMMEDIATE EARLY GLYCOPROTEIN PRECURSOR.
GN UL37.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88265854.
RA Kouzarides T., Bankier A.T., Satchwell S.C., Preddy E., Barrell B.G.;
RT "An immediate early gene of human cytomegalovirus encodes a potential
membrane glycoprotein."
RL Virology 165:151-164(1988).
RN [2]
RP COMPLETE GENOME.
RX MEDLINE; 90269039.
RA Chee M.S., Bankier A.T., Bohni R., Brown C.M., Cerny R.,
RA Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,
RA Preddy E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
cytomegalovirus strain AD169."
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X17403; CAA35396.1; -
DR PIR; S09801; Q0BEU5.
KW Glycoprotein; Transmembrane; Signal.

FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 487 IMMEDIATE EARLY GLYCOPROTEIN.
FT TRANSMEM 433 459 POTENTIAL.
FT CARBOHYD 206 206 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 210 210 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 219 219 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 223 223 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 281 281 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 337 337 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 343 343 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 379 379 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 391 391 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 487 AA; 56124 MW; FA39A37C51FFB4A3 CRC64;

Query Match 28.1%; Score 49.5; DB 1; Length 487;
Best Local Similarity 38.7%; Pred. No. 22;
Matches 12; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

QY 5 SNKEDSGRRDDILMEKPSRPMSNP-DTEG 34
DB 91 TEREDTEEREDEEEENEARTPEVNPIDAEG 121

RESULT 9
POLG_FMDV1
ID POLG_FMDV1 STANDARD; PRT; 2333 AA.
AC Q03306; Q84750; Q84751; Q84752; Q84753; Q84754; Q89824; Q84760;
AC Q84761; Q84762; Q84763; Q84764; Q84765; Q84766; Q84767; Q64768;
AC Q84768; Q84769;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: NONSTRUCTURAL PROTEIN P20A; COAT
PROTEINS VP1 TO VP4; CORE PROTEIN P52; GENOME-LINKED PROTEINS VP61 TO
DE VP63; PICORNAIN 3C (EC 3.4.22.28) (PROTEASE 3C) (P3C); RNA-DIRECTED
DE RNA POLYMERASE P56A (EC 2.7.7.48)].
OS Foot-and-mouth disease virus (strain A10-61) (Aphthovirus A).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Aphthovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 84169547.
RA Carroll A.R., Rowlands D.J., Clarke B.E.;
RT "The complete nucleotide sequence of the RNA coding for the primary
translation product of foot and mouth disease virus."
RL Nucleic Acids Res. 12:2461-2472(1984).
RN [2]
RP SEQUENCE OF 115-1048 FROM N.A.
RX MEDLINE; 82211814.
RA Boothroyd J.C., Harris T.J.R., Rowlands D.J., Lowe P.A.;
RT "The nucleotide sequence of cDNA coding for the structural proteins
of foot-and-mouth disease virus."
RL Gene 17:153-161(1982).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>), or send an email to license@isb-sib.ch.

```

CC -----
CC EMBL; X74812; CAA52812.1; -.
CC HSSP; P13899; 1TWE.
CC INTERPRO; IPR000605; -.
CC INTERPRO; IPR001205; -.
CC INTERPRO; IPR001676; -.
CC PFAM; PF00680; RNA_dep_RNA_pol; 1.
CC PFAM; PF00910; RNA_helicase; 1.
CC PFAM; PF00073; rhv; 5.
CC Polyprotein; Coat protein; RNA-directed RNA polymerase;
CC Transferase; Hydrolase; Thiol protease; Nonstructural protein;
CC Myristate.
CC CHAIN 1 201 NONSTRUCTURAL PROTEIN P20A.
CC CHAIN 202 286 COAT PROTEIN VP4.
CC CHAIN 287 504 COAT PROTEIN VP2.
CC CHAIN 505 724 COAT PROTEIN VP3.
CC CHAIN 725 938 COAT PROTEIN VP1.
CC CHAIN 939 954 CORE PROTEIN X.
CC CHAIN 955 1108 CORE PROTEIN P14.
CC CHAIN 1109 1426 CORE PROTEIN P41.
CC CHAIN 1427 1579 CORE PROTEIN P19.
CC CHAIN 1580 1602 GENOME-LINKED PROTEIN VP61.
CC CHAIN 1603 1826 GENOME-LINKED PROTEIN VP62.
CC CHAIN 1827 1850 GENOME-LINKED PROTEIN VP63.
CC CHAIN 1851 1863 PROTEASE.
CC CHAIN 1864 2336 RNA-DIRECTED RNA POLYMERASE.
CC CHAIN 202 202 MYRISTATE (BY SIMILARITY).
CC LIPID 202 202
CC SEQUENCE 2336 AA; 259982 MW; 15AC2AB022B5B954 CRC64;

Query Match 30.1%; Score 53; DB 1; Length 2336;
Best Local Similarity 37.5%; Pred. No. 46;
Matches 9; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 11 GQRDDILMKRSPRMESNPDEG 34
Db 1557 GHRASDDVNSEPARVPEQPAEG 1580

RESULT 6
MUTB_STRCM
ID MUTB_STRCM STANDARD; PRT; 733 AA.
AC Q05065;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE METHYLMALONYL-COA MUTASE ALPHA-SUBUNIT (EC 5.4.99.2) (MCM-ALPHA).
GN MUTB.
OS Streptomyces cinnamonensis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3823.5;
RA MEDLINE; 93273720.
RA Birch A., Leiser A., Robinson J.A.;
RT "Cloning, sequencing, and expression of the gene encoding
RT methylmalonyl-coenzyme A mutase from Streptomyces cinnamonensis.";
RL J. Bacteriol. 175:3511-3519(1993).
CC -!- FUNCTION: CATALYZES THE ISOMERIZATION OF SUCCINYL-COA TO
CC METHYLMALONYL-COA DURING SYNTHESIS OF PROPIONATE FROM
CC TRICARBOXYLIC ACID-CYCLE INTERMEDIATES. THIS CONVERSION MOST
CC LIKELY REPRESENTS AN IMPORTANT SOURCE OF BUILDING BLOCKS FOR
CC POLYKETIDE ANTIBIOTIC BIOSYNTHESIS. IT IS UNABLE TO CATALYZE
CC THE CONVERSION OF ISOBUTYRYL-COA INTO N-BUTYRYL-COA.
CC -!- CATALYTIC ACTIVITY: (R)-2-METHYL-3-OXOPROPYL-COA =
CC SUCCINYL-COA.
CC -!- COFACTOR: ADENOSYLCOBALAMIN.
CC -!- PATHWAY: PROPIONIC ACID FERMENTATION.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -!- SIMILARITY: BELONGS TO THE METHYLMALONYL-COA MUTASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/),
or send an email to license@isb-sib.ch).
-----
CC EMBL; L10064; AAA03041.1; -.
CC PIR; B40595; B40595.
CC HSSP; P11653; 3REQ.
CC INTERPRO; IPR001629; -.
CC PFAM; PF01642; MM_CoA_mutase; 1.
CC PROSITE; PS00544; METMALONYL_COA_MUTASE; 1.
CC Isomerase; Vitamin B12; Cobalt.
KW DOMAIN 571 610 METHYLMALONYL-COA-BINDING SITE (BY
FT SIMILARITY).
FT SEQUENCE 733 AA; 79454 MW; F02EEB54A154FA59 CRC64;

Query Match 29.5%; Score 52; DB 1; Length 733;
Best Local Similarity 34.4%; Pred. No. 16;
Matches 11; Conservative 5; Mismatches 12; Indels 4; Gaps 1;

QY 7 KEDSGRQDDILMKRPS----RPMESNPDEG 34
Db 29 KESVGKSESDLLWETPEGIAVKPLYTGADVEG 60

RESULT 7
P5CS_ORYSA
ID P5CS_ORYSA STANDARD; PRT; 716 AA.
AC O04226;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DELTA 1-PYRROLINE-5-CARBOXYLATE SYNTHETASE (P5CS) [INCLUDES: GLUTAMATE
DE 5-KINASE (EC 2.7.2.11) (GAMMA-GLUTAMYL KINASE) (GK); GAMMA-GLUTAMYL
DE PHOSPHATE REDUCTASE (GPR) (EC 1.2.1.41) (GLUTAMATE-5-SEMIALDEHYDE
DE DEHYDROGENASE) (GLUTAMYL-GAMMA-SEMIALDEHYDE DEHYDROGENASE)].
GN P5CS.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. AKIBARE;
RX MEDLINE; 97260389.
RA Igarashi Y., Yoshida Y., Sanada Y., Yamaguchi-Shinozaki K., Wada K.,
RA Shinozaki K.;
RT "Characterization of the gene for delta1-pyrroline-5-carboxylate
RT synthetase and correlation between the expression of the gene and
RT salt tolerance in Oryza sativa L.;"
RL Plant Mol. Biol. 33:857-865(1997).
CC -!- FUNCTION: P5CS PLAYS A KEY ROLE IN PROLINE BIOSYNTHESIS, LEADING
CC TO OSMOREGULATION IN PLANTS.
CC -!- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE = ADP + L-GLUTAMATE
CC 5-PHOSPHATE (PRODUCT RAPIDLY CYCLIZES TO 5-OXOPROLINE AND
CC ORTHOPHOSPHATE).
CC -!- CATALYTIC ACTIVITY: L-GLUTAMATE 5-SEMIALDEHYDE + ORTHOPHOSPHATE +
CC NADP(+) = L-GAMMA-GLUTAMYL 5-PHOSPHATE + NADPH.
CC -!- ENZYME REGULATION: FEEDBACK REGULATED BY PROLINE.
CC -!- PATHWAY: FIRST AND SECOND STEPS IN PROLINE BIOSYNTHESIS PATHWAY.
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LEAVES AND IS
CC INDUCIBLE IN ROOTS SUBJECTED TO SALT STRESS.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE GLUTAMATE 5-
CC KINASE FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE GAMMA-
CC GLUTAMYL PHOSPHATE REDUCTASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 09:03:58 ; Search time 58.45 Seconds
(without alignments)
18.584 Million cell updates/sec

Title: US-09-236-468A-2_COPY_508_541

Perfect score: 176

Sequence: 1 LSTRSKEDSGRDDILMEKPSRPMESNPDTTEG 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 87993

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	%	Description
1	141	80.1	550	1	PTB2_HUMAN
2	59	33.5	546	1	PTH2_RAT
3	54.5	31.0	1589	1	CS25_YEAST
4	53	30.1	451	1	GAGY_DROME
5	53	30.1	2336	1	POLG_FMDVZ
6	52	29.5	733	1	MUTB_STRCM
7	51	29.0	716	1	P5CS_ORYSA
8	49.5	28.1	487	1	VGLI_HCMVA
9	49.5	28.1	2333	1	POLG_FMDV1
10	49	27.8	240	1	HDGF_HUMAN
11	49	27.8	2039	1	CCH1_YEAST
12	48.5	27.6	640	1	ELM1_YEAST
13	47.5	27.0	237	1	HDGF_MOUSE
14	47.5	27.0	712	1	SSK1_YEAST
15	47.5	27.0	1082	1	SP23_YEAST
16	47.5	27.0	3924	1	ANK2_HUMAN
17	47	26.7	317	1	YM91_YEAST
18	47	26.7	350	1	PEX1_YEAST
19	47	26.7	388	1	ZNT3_MOUSE
20	47	26.7	446	1	KCCL1_YEAST
21	47	26.7	460	1	SYS_HALMA
22	47	26.7	586	1	ESR1_XENLA
23	47	26.7	2332	1	POLG_FMDVO
24	46.5	26.4	518	1	YLW6_CAEEL
25	46.5	26.4	726	1	P5C2_ARATH
26	46	26.1	238	1	VGLG_HSV11
27	46	26.1	354	1	PHN1_ECOLI
28	46	26.1	826	1	SSP2_PLAYO
29	46	26.1	992	1	EBN6_ERV
30	46	26.1	1018	1	KY26_CAEEL
31	45.5	25.9	453	1	SSF2_YEAST
32	45.5	25.9	509	1	NFIL_RAT
33	45.5	25.9	522	1	NFIA_CHICK

34	45.5	25.9	976	1	XPR6_YARLI	P42781 yarrowia li
35	45	25.6	146	1	PAG1_HUMAN	O75459 homo sapien
36	45	25.6	307	1	IF2B_MALDO	P55871 malus domes
37	45	25.6	397	1	CSP_PLAFO	P19597 plasmodium
38	45	25.6	407	1	Z174_HUMAN	Q15697 homo sapien
39	45	25.6	513	1	DMP1_HUMAN	Q13316 homo sapien
40	45	25.6	524	1	SBP_SOYBN	Q04672 glycine max
41	45	25.6	551	1	CEA6_ECOLI	P17999 escherichia
42	45	25.6	892	1	BNI4_YEAST	P53858 saccharomyc
43	45	25.6	1062	1	CC7_SCHPO	P41892 schizosacch
44	45	25.6	1093	1	TMF1_HUMAN	P82094 homo sapien
45	45	25.6	1220	1	IF2P_HUMAN	O60841 homo sapien

ALIGNMENTS

RESULT 1

ID	PTR2_HUMAN	STANDARD;	PRT;	550 AA.
AC	P49190;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	PARATHYROID HORMONE RECEPTOR PRECURSOR (PTH2 RECEPTOR).			
GN	PTH2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.			
RC	TISSUE=BRIN;			
RX	MEDLINE: 95318121.			
RA	Usdin T.B., Gruber C., Bonner T.I.;			
RT	"Identification and functional expression of a receptor selectively recognizing parathyroid hormone, the PTH2 receptor.";			
RL	J. Biol. Chem. 270:15455-15458(1995).			
RN	[2]			
RP	SEQUENCE OF 26-40 AND 306-550 FROM N.A.			
RX	MEDLINE: 97079671.			
RA	Usdin T.B., Modi W., Bonner T.I.;			
RT	"Assignment of the human PTH2 receptor gene (PTH2) to chromosome 2q33 by fluorescence in situ hybridization.";			
RL	Genomics 37:140-141(1996).			
CC	!- FUNCTION: THIS IS A SPECIFIC RECEPTOR FOR PARATHYROID HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYL CYCLASE.			
CC	!- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN BRAIN AND PANCREAS.			
CC	ALSO EXPRESSED IN THE TESTIS.			
CC	!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL: U25128; AAC50157.1; -			
DR	EMBL: U47124; AAC96796.1; -			
DR	EMBL: U47129; AAC50767.1; -			
DR	EMBL: U47125; AAC50767.1; JOINED.			
DR	EMBL: U47126; AAC50767.1; JOINED.			
DR	EMBL: U47127; AAC50767.1; JOINED.			
DR	EMBL: U47128; AAC50767.1; JOINED.			
DR	CCRD: GCR_2003; -			
DR	MIM: 601469; -			
DR	INTERPRO: IPR000832; -			
DR	PFAM: PF00002; 7tm2; 1			
DR	PRINTS: PRO0249; GPCSECRETIN.			
DR	PROSITE: PS00649; G_PROTEIN_RECEP_F2_1; 1.			
DR	PROSITE: PS00650; G_PROTEIN_RECEP_F2_2; 1.			

THIS PAGE BLANK (USPTO)

Db 143 RTKKKNSGRQFLRYNEKTPSEPMEKLP 169

RESULT 15

H69125

hypothetical protein MTH210 - Methanobacterium thermoautotrophicum (strain Delta H)

C:Species: *Methanobacterium thermoautotrophicum*

C;Species: mechanobacterium thermophilum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C;Accession: H69125

R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Gibson, N.

ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A;Title: Complete genome sequence of Meth

A; Reference number:

A;Accession: H69125

A;Status: preliminary

A; Molecule type: DNA

A:Residues: 1-520 <MTH>

A; Cross-references: GB: AE000808; GB: A

A; Experiment

C;Genetics:
A:Conc: MTH210

Query Match	28.48	Score 50:	DB 2:	Length 520:
-------------	-------	-----------	-------	-------------

Query Muccn	28.4%	Score 50; DB
Best Local Similarity	38.5%	Pred. No. 39;

Best local similarity 50.54, PFD: NO. 33,
Matches 10; Conservative 8; Mismatches 6; Indels 2; Gaps 1;

QY 8 EDGRQRDD--ILMEKPSRPMESNPD 31

[illegible]

Db 273 DD LGRQDDHEVILNRLIVPLENSKD 298

Search completed: November 8, 2000, 08:53:39

Job time: 370 sec

A:Accession: T06055
A:Molecule type: DNA
A:Residues: 1-834 <BEV>
A:Cross-references: EMBL:AL035679; GSPDB:GN00062; ATSP:F19H22.50
A:Experimental source: cultivar Columbia; BAC clone F19H22
C:Genetics:
A:Gene: ATSP:F19H22.50
A:Map position: 4
A:Introns: 63/1; 126/3; 196/2; 234/3; 276/2; 389/3; 422/3; 633/3; 700/3
C:Superfamily: kinesin heavy chain; kinesin motor domain homology
F:14-339/Domain: kinesin motor domain homology <KNOT>

Query Match 29.5%; Score 52; DB 2; Length 834;
Best Local Similarity 47.8%; Pred. No. 36;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 5 SNKEDSGRQRDDILMEKPSRPME 27
 || ||||| : | : | : |
Db 635 SNAEDSGMKRRDIGDSMDREFE 657

RESULT 9
H70863
hypothetical protein Rv2449c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: H70863
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, K.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajadream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Soares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987
A:Accession: H70863
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-419 <COL>
A:Cross-references: GB:AL021246; GB:AL123456; MID:g3261507; PIDN:CAAL6026.1; PID:g279
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv2449c

Query Match 29.0%; Score 51; DB 2; Length 419;
Best Local Similarity 48.1%; Pred. No. 22;
Matches 13; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

Qy 1 LSTRSNKEDSGRQRDDILMEKPSRPME 27
 ||||| : | : | : | : |
Db 192 LSTASNDPDRRLSDPYMLSPDRGAE 218

RESULT 10
T25639
hypothetical protein C46H11.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T25639
R:Miller, N.; Bradshaw, H.; Wamsley, P.
submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans cosmid C46H11.
A:Reference number: Z20061
A:Accession: T25639
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-486 <MIL>
A:Cross-references: EMBL:U88314; PIDN:AB42358.1; GSPDB:GN00019; CESP:C46H11.4
A:Experimental source: strain Bristol N2; clone C46H11
C:Genetics:
A:Gene: CESP:C46H11.4
A:Map position: 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: November 8, 2000, 08:53:37 ; Search time 99.87 Seconds
(without alignments)
21.605 Million cell updates/sec

Title: US-09-236-468A-2_COPY_508_541
Perfect score: 176
Sequence: 1 LSTRSNKEDSGRQDDILMEKPSRPMSNPDTG 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues
Total number of hits satisfying chosen parameters: 182106

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	141	80.1	550 2 A57519	parathyroid hormon
2	54.5	31.0	1589 1 RBYC5	cell division cont
3	54	30.7	1390 2 T18883	hypothetical prote
4	53	30.1	451 1 FOFGY	retrovirus-related
5	53	30.1	611 2 T06458	nucleolin homolog
6	53	30.1	2336 2 S37077	genome polypeptide
7	52	29.5	733 2 B40595	methylmalonyl-CoA
8	52	29.5	834 2 T06055	hypothetical prote
9	51	29.0	419 2 H70863	hypothetical prote
10	51	29.0	486 2 T25639	hypothetical prote
11	51	29.0	494 2 T42444	1D-myo-inositol-tr
12	51	29.0	716 2 T03695	delta 1 pyroline-
13	50	28.4	490 2 T25430	hypothetical prote
14	50	28.4	510 2 S10724	surface protein pr
15	50	28.4	520 2 H69125	hypothetical prote
16	50	28.4	1214 2 T21915	hypothetical prote
17	49.5	28.1	487 1 Q0BEU5	UL37 protein precu
18	49.5	28.1	2333 1 GNNY2F	genome polypeptide
19	49	27.8	240 2 A55055	hepatoma-derived g
20	49	27.8	416 2 T22073	hypothetical prote
21	49	27.8	1025 2 S69790	fibronectin-bindin
22	49	27.8	1225 2 T18954	hypothetical prote
23	49	27.6	2039 2 S64540	probable calcium c
24	48.5	27.6	640 1 S37869	probable serine/th
25	48	27.3	250 2 C71611	probable secreted
26	48	27.3	454 2 T45297	tyrosine phenol-ly
27	48	27.3	887 2 T20866	hypothetical prote
28	48	27.3	1229 2 T48959	kinesin-like prote
29	48	27.3	1282 2 JC4393	microbial collagen

ALIGNMENTS

RESULT 1

A57519
parathyroid hormone receptor 2 precursor - human
N:Alternate names: PTH2 receptor
C:Species: Homo sapiens (man)
C:Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 24-Nov-1999
C:Accession: A57519
R:Usdin, T.B.; Gruber, C.; Bonner, T.I.
J. Biol. Chem. 270, 15455-15458, 1995
A:Title: Identification and functional expression of a receptor selectively recogni
A:Reference number: A57519; MUID:95318121
A:Accession: A57519
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-550 <USD>
A:Cross-references: GB:U25128; MID:g887966; PIDN:AAC50157.1; PID:g887967
C:Genetics:
A:Gene: GDB:PTH2; PTHR2
A:Cross-references: GDB:731977; OMIM:601469
A:Map position: 2q33-2q33
C:Superfamily: glucagon receptor
C:Keywords: hormone receptor

Query Match 80.1%; Score 141; DB 2; Length 550;
Best Local Similarity 96.4%; Pred. No. 1.7e-11;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 KEDSGRQDDILMEKPSRPMSNPDTG 34

Db 514 KEDSGRQDDILMEKPSRPMSNPDTG 541

RESULT 2

RBYC5
cell division control protein CDC25 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein L2142.6; protein YLR310c
C:Species: Saccharomyces cerevisiae
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
C:Accession: A26596; S51442; A23444; S43051; S47990
R:Broek, D.; Toda, T.; Michaeli, T.; Levin, L.; Birchmeier, C.; Zoller, M.; Powers, S
Cell 48, 789-799, 1987
A:Title: The S. cerevisiae CDC25 gene product regulates the RAS/adenylate cyclase pat
A:Reference number: A26596; MUID:87131091
A:Accession: A26596
A:Molecule type: DNA
A:Residues: 1-1589 <BRO>
A:Cross-references: EMBL:M15458; NID:g171184; PIDN:AAA34478.1; PID:g171185
R:Pauley, A.
submitted to the EMBL Data Library, November 1994
A:Description: The sequence of S. cerevisiae cosmid L2142.
A:Reference number: S51437

THIS PAGE BLANK (USPTO)

```
QY      13 QRDDILMEKPSRPMESNPDTE 33
         | | : | | | | : | | : :
Db      104 QSDELLHSKSHSPLDSNQTS 124
```

Search completed: November 8, 2000, 08:49:25
Job time: 118 sec


```

RESULT" 11
US-08-760-745-5
: Sequence 5, Application US/08760745
: Patent No. 5972658
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Goli, Surya K.
: APPLICANT: Muri, Lynn E.
: TITLE OF INVENTION: LUNG GROWTH FACTOR VA
: NUMBER OF SEQUENCES: 5
: ADDRESS: INCYTE PHARMACEUTICALS, INC.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: US
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS

```

ADDRESSES: A. David Spevack
STREET: NMRC Building 1 T-12 National Naval
STREET: Medical Center
CITY: Bethesda
STATE: MD
COUNTRY: USA
ZIP: 20814-5044
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/638,431
FILING DATE: 19910110
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Spevack, Avrom D.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 295-6759
TELEFAX: (301) 295-4033
INFORMATION FOR SEQ ID NO: 2:


```

RESULT 7
US-08-760-745-1
; Sequence 1, Application US/08760745
; Patent No. 5972658
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Svyra K.
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: LUNG GROWTH FACTOR VARIANT
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760.745
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

Db 508 LSTRNKEDSGRQDDILMEKPSRPMSNPDTG 541

RESULT 2

PCT-US95-07085-2
; Sequence 2, Application PC/TUS9507085
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone Receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07085
; FILING DATE: 05-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-393
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-07085-2

Query Match 100.0%; Score 176; DB 4; Length 541;
Best Local Similarity 100.0%; Pred. No. 3.7e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSTRNKEDSGRQDDILMEKPSRPMSNPDTG 34
|||||

Db 508 LSTRNKEDSGRQDDILMEKPSRPMSNPDTG 541
|||||

RESULT 3

US-08-468-557-2
; Sequence 2, Application US/0846857
; Patent No. 5759833
; GENERAL INFORMATION:
; APPLICANT: Shiba, Kiyotaka
; APPLICANT: Kranz, Janice E.
; APPLICANT: Schimmel, Paul R.
; TITLE OF INVENTION: Human Isoleucyl-tRNA Synthetase
; TITLE OF INVENTION: Proteins, Nucleic Acids and Tester Strains Comprising Same
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US

; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,557
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/250,852
; FILING DATE: 27-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: CPI94-13A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 993 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-557-2

Query Match 34.1%; Score 60; DB 1; Length 993;
Best Local Similarity 38.9%; Pred. No. 1.6;
Matches 14; Conservative 5; Mismatches 13; Indels 4; Gaps 1;

Qy 2 STRSNKEDSGRQDDILMEKPSRPME----SNPDTG 33
|||||

Db 30 SNHQPNNSGRYRDTVLLPQTSPFKLLGRQOPDTG 65
|||||

RESULT 4

US-09-356-952-4
; Sequence 4, Application US/09356952
; Patent No. 6117663
; GENERAL INFORMATION:
; APPLICANT: Borlack-Sjodin, Ann
; APPLICANT: Margarit, S. M.
; APPLICANT: Bor-Sogli, Dafna
; APPLICANT: Cole, Philip
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
; FILE REFERENCE: 600-1-228N
; CURRENT APPLICATION NUMBER: US/09/356,952
; CURRENT FILING DATE: 1999-07-19
; EARLIER APPLICATION NUMBER: 60/093,631
; EARLIER FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1589
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-09-356-952-4

Query Match 31.0%; Score 54.5; DB 3; Length 1589;
Best Local Similarity 36.4%; Pred. No. 16;
Matches 12; Conservative 8; Mismatches 8; Indels 5; Gaps 1;

Qy 2 STRSNKEDSGRQDDILMEKPSRPMSNPDTG 34
|||||

Db 365 STRSKQSESEQNRSLLMEK-----QDSKETDG 392
|||||

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:49:24 ; Search time 97.15 Seconds
(without alignments)
5.866 Million cell updates/sec

Title: US-09-236-468A-2_COPY_508_541

Perfect score: 176
Sequence: 1 LSTRSNKEDSGRQDDILMEKPSRPMESNPDTG 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 164575

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6_COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/PTCUS_COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	176	100.0	541	3	US-08-468-011A-2
2	176	100.0	541	4	PCT-US95-07085-2
3	60	34.1	993	1	US-08-468-557-2
4	54.5	31.0	1589	3	US-08-356-952-4
5	49	27.8	39	1	US-08-294-189-6
6	49	27.8	120	1	US-08-294-189-5
7	49	27.8	235	2	US-08-760-745-1
8	49	27.8	240	2	US-08-760-745-3
9	48.5	27.6	563	2	US-08-061-636-2
10	48.5	27.6	563	4	PCT-US94-05268-2
11	47.5	27.0	237	2	US-08-760-745-5
12	46	26.1	826	1	US-07-638-431-2
13	46	26.1	826	4	PCT-US92-00018-2
14	45	25.6	59	3	US-08-812-121-12
15	45	25.6	170	1	US-08-818-163A-2
16	45	25.6	1093	4	PCT-US93-03077-1
17	44	25.0	38	1	US-08-340-428B-5
18	44	25.0	38	4	PCT-US93-07306-5
19	44	25.0	439	3	US-08-921-209-2
20	44	25.0	439	3	US-09-411-763-2
21	44	25.0	451	3	US-08-996-139-4
22	44	25.0	532	2	US-08-579-777A-2
23	44	25.0	591	3	US-08-996-139-2
24	44	25.0	596	2	US-08-836-620A-16
25	44	25.0	616	3	US-08-996-139-6
26	44	25.0	1588	1	US-08-698-551-16
27	44	25.0	1588	2	US-08-602-228-16
28	44	25.0	1588	2	US-08-839-032A-16

29	44	25.0	2482	1	US-08-338-254-6	Sequence 6, Appli
30	43	24.4	167	1	US-07-688-352C-26	Sequence 26, Appl
31	43	24.4	167	2	US-08-474-379C-26	Sequence 26, Appl
32	43	24.4	167	3	US-09-146-249A-26	Sequence 26, Appl
33	43	24.4	167	3	US-08-206-188B-26	Sequence 26, Appl
34	43	24.4	250	1	US-08-144-121-7	Sequence 7, Appli
35	43	24.4	250	2	US-08-460-309-7	Sequence 7, Appli
36	43	24.4	250	2	US-08-125-077-7	Sequence 7, Appli
37	43	24.4	250	2	US-08-735-893-7	Sequence 7, Appli
38	43	24.4	363	3	US-08-651-999A-4	Sequence 4, Appli
39	43	24.4	746	1	US-08-476-519-11	Sequence 11, Appl
40	43	24.4	746	4	PCT-US95-09323-11	Sequence 3, Appli
41	43	24.4	752	3	US-08-709-784-3	Sequence 9, Appli
42	43	24.4	756	2	US-08-209-521-9	Sequence 12, Appl
43	43	24.4	756	2	US-08-209-521-12	Sequence 27, Appl
44	43	24.4	756	2	US-08-209-521-27	Sequence 2, Appli
45	43	24.4	777	1	US-08-476-519-2	

ALIGNMENTS

RESULT 1
US-08-468-011A-2
; Sequence 2, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESS: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08468.011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-011A-2

Query Match 100.0%; Score 176; DB 3; Length 541;
Best Local Similarity 100.0%; Pred. No. 3.7e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 LSTRSNKEDSGRQDDILMEKPSRPMESNPDTG 34
|||||

R27118
ID R27118 standard; Protein; 321 AA.
XX
AC R27118;
XX
DT 01-MAR-1993 (first entry)
XX
DE Fucose dehydrogenase.
XX
KW Arthrobacter oxidans; F1; induction; assay.
XX
OS Arthrobacter oxidans F1.
XX
PN EP506262-A.
XX
PD 30-SEP-1992.
XX
PF 13-MAR-1992; 92EP-0302170.
XX
PR 29-MAR-1991; 91JP-0089184.
XX
PA (TAKI) TAKARA SHUZO CO LTD.
XX
PI Kato I, Kotani H, Mitta M, Sakai T;
XX
DR WPI; 1992-325548/40.
DR P-PDSB; R27118.
XX
PT Isolated gene encoding L-fucose dehydrogenase - useful for prodn.
PT of enzyme by genetic engineering
XX
PS Disclosure; Page 8; 16pp; English.
XX
CC Genomic DNA from Arthrobacter oxidans F1 was subjected to
CC restriction enzyme analysis and the N-terminal amino acid sequence
CC of L-fucose dehydrogenase determined. A degenerate probe was
CC synthesised based on this amino acid sequence. The probe was used
CC to screen an Arthrobacter cDNA library to isolate a L-fucose dehydro-
CC genase clone. The isolation of such a clone provides a convenient
CC method for prodn. of L-fucose dehydrogenase without the need for
CC induction by L-fucose. The probe may be used to evaluate the extent
CC of expression of L-fucose dehydrogenase. The DNA sequence is
CC widely used to assay L-fucose levels. The protein sequence of
CC L-fucose dehydrogenase may be used to produce antibodies.
XX
SQ Sequence 321 AA;

Query Match 26.1%; Score 46; DB 13; Length 321;
Best Local Similarity 41.9%; Pred. No. 90;
Matches 13; Conservative 4; Mismatches 8; Indels 6; Gaps 2;

Qy 10 SGRQRDD-ILMEKPSRPMSNP-----DTEG 34
:|||||:|:|:|
Db 70 agqgrdsylvlstkigrllrpnpspgkdtg 100

Search completed: November 8, 2000, 08:51:53
Job time: 265 sec

XX FH Key Location/Qualifiers
 FT Peptide 259..264
 FT Peptide /note= "Ser/Thr kinase consensus sequence"
 FT Peptide 309..317
 FT Peptide /note= "Ser/Thr kinase consensus sequence"
 XX W09426885-A.

XX PD 24-NOV-1994.

XX PD 12-MAY-1994; 94WO-US05268.

XX PD 12-MAY-1993; 93US-0061636.

XX PD (INRM) INST NAT SANTE & RECH MEDICALE.

XX PD (INSP) INST PASTEUR.

XX PD (IOWA) UNIV IOWA STATE RES FOUND INC.

XX PI Madaule P, Myers AM;

XX DR WPI; 1995-006782/01.

XX DR N-PSDB; Q75120.

XX XX Saccharomyces cerevisiae strains contg. a pseudohyphal growth

XX PT mutant gene and exhibiting constitutive pseudohyphal growth

XX PT used to control pathogenic transformation in fungi, and to

XX PT degrade polysaccharide(s) in fermentation processes

XX PS Disclosure; Page 31-35; 59pp; English.

XX XX S. cerevisiae wild-type gene ELM1 (Elongated Morphology) codes for a

XX CC novel protein-kinase homolog required for yeast morphology.

XX CC Deletion or mutation of ELM1 causes constitutive pseudohyphal growth.

XX CC Manipulation of the gene can be used to control pathogenic

XX CC transformation in fungi, to control plant pathogens, or to produce

XX CC yeasts capable of fermenting polysaccharides to ethanol.

XX XX Sequence 563 AA;

XX SQ

XX Query Match 27.6%; Score 48.5; DB 16; Length 563;

XX Best Local Similarity 28.9%; Pred. No. 79;

XX Matches 11; Conservative 9; Mismatches 13; Indels 5; Gaps 1;

XX Qy 1 LSTRSKEDSGRQRDILME-----KPSRPMSNPDT 33

XX Db 442 ltkgkktsgkdkvlvsatskvtptsihiideepdke 479

XX RESULT 11

XX Y13459

XX ID Y13459 standard; protein; 730 AA.

XX AC Y13459;

XX XX 26-JUL-1999 (first entry)

XX DE Amino acid sequence of human Fe65-like protein.

XX KW Phosphotyrosine interaction domain; PI domain; modulator; APP;

XX KW amyloid precursor protein; cytoplasmic domain; Alzheimer's disease;

XX KW neurodegenerative disease; dementia; Fe65.

XX OS Homo sapiens.

XX XX W09921995-A1.

XX PD 06-MAY-1999.

XX XX 23-OCT-1998; 98WO-US22523.

XX PF 17-APR-1998; 98US-0062085.

XX PR

PR 24-OCT-1997; 97US-0957660.

XX XX (UYRQ) UNIV ROCKEFELLER.

XX PI Buxbaum J, Greengard P, Sabo S;

XX XX WPI; 1999-326703/27.

XX DR Cell line for screening agents that inhibit processing of amyloid

XX PT precursor protein

XX XX Claim 4; Fig 14; 86pp; English.

XX XX The invention relates to modulators that affect the interaction of a

XX CC polypeptide defining a PI (phosphotyrosine interaction) domain, or its

XX CC conserved variants or fragments, with APP (amyloid precursor protein).

XX CC The polypeptide defining a PI domain binds to the cytoplasmic domain of

XX CC APP, regulating the secretion of APP fragments. By modulating this

XX CC interaction, the modulator can inhibit or prevent APP processing and

XX CC trafficking, and thus progress or onset of Alzheimer's disease. The

XX CC modulators are used to treat and/or prevent neurodegenerative diseases

XX CC in mammals, particularly Alzheimer's disease and dementia. Sequences

XX CC Y1348-461 represent examples of polypeptides defining a PI domain.

XX XX Sequence 730 AA;

XX SQ

XX Query Match 27.3%; Score 48; DB 20; Length 730;

XX Best Local Similarity 50.0%; Pred. No. 1.3e+02;

XX Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

XX Qy 16 DILMEKPSRPMSNP 31

XX Db 177 dlilqknpqsped 192

XX RESULT 12

XX W33216

XX ID W33216 standard; Protein; 237 AA.

XX AC W33216;

XX XX 27-APR-1998 (first entry)

XX DT Mouse hepatoma derived growth factor.

XX DE Hepatoma derived growth factor; HDGF; mouse.

XX KW Mus sp.

XX OS JP09252777-A.

XX PN 30-SEP-1997.

XX PD 21-MAR-1996; 96JP-0064001.

XX PF 21-MAR-1996; 96JP-0064001.

XX PR (KISH/) KISHIMOTO C.

XX XX (SEKI) SEKISUI CHEM IND CO LTD.

XX PA WPI; 1997-530150/49.

XX DR N-PSDB; T88420, V03867.

XX XX DNA encoding mouse hepatoma derived growth factor - and transformed

XX PT E. coli or animal cell for its recombinant production

XX XX Claim 1; Page 8-9; 13pp; Japanese.

XX PS This sequence represents the mouse hepatoma derived growth factor (HDGF)

XX CC and is encoded by the DNA molecule of the invention. The DNA encoding

XX CC this sequence can be used in a vector which is used to transform an

XX CC E. coli or animal cell. The transformed cell can be used for the

```

PF 02-AUG-1999; 99WO-US17545.
XX
PR 10-AUG-1998; 98US-0132028.
PA (SMK ) SMITHLINE BEECHAM CORP.
XX
PI Wilding EI, Black MT, Traini CM;
XX
DR WPI; 2000-224274/19.
DR N-PSDB; 261497.
XX
XX New nrdE polypeptide from Staphylococcus aureus, useful e.g. for
PT vaccination against bacterial infection and for drug screening -
PT
PS Claim 1; Page 59-61; 64pp; English.
XX
XX The present sequence represents a nrdE polypeptide. The polypeptide
CC is used to screen for specific agonists and antagonists; to treat
CC conditions that require increased activity or expression of nrdE; to
CC raise specific antibodies; to identify receptors; and in vaccines.
CC The polynucleotide is used for recombinant (or in vivo) production
CC of the nrdE polypeptide, and as sources of antisense sequences that
CC inhibit expression, or of probes and primers. Detecting mutations
CC in nrdE-encoding genomic sequences, or measuring the expression of
CC nrdE, can be used for diagnosis, staging and prognosis of disease (or
CC susceptibility), also for serotyping or chromosome identification.
CC Diseases which may be diagnosed or treated are particularly infection
CC by S. aureus, but may also be infection by Helicobacter pylori, and
CC associated ulcers and cancers.
XX
SQ Sequence 701 AA;

Query Match 27.8%; Score 49; DB 21; Length 701;
Best Local Similarity 44.4%; Pred. No. 87;
Matches 12; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 7 KEDSGQRDDILMEKPSRPMSNPDTPE 33
Dl 308 keeygtldldlekyddmvanpnve 334

RESULT 9
ID R70126
XX R70126 standard; Protein; 1025 AA.
AC R70126;
XX
DT 31-OCT-1995 (first entry)
XX
DE Serum opacity factor (SOF22).
XX
XX Serum opacity factor; diagnostic; high density lipoprotein; HDL;
KW apolipoprotein; coronary disease; Streptococcus pyogenes.
XX
XX Streptococcus pyogenes strain 22 D734.
XX
XX Key Location/Qualifiers
FH Peptide 1..29
FT /note= "signal sequence"
FT Protein 30..1025
FT /note= "mature SOF22"
FT Cleavage-site 29..30
FT /note= "putative signal sequence cleavage site"
FT Region 780..807
FT /note= "proline-rich motif present in the majority
FT of streptococcal surface proteins"
FT
FT Region 820..828
FT /note= "proline-rich motif present in the majority
FT of streptococcal surface proteins"
FT
FT Region 849..857
FT /note= "proline-rich motif present in the majority
FT of streptococcal surface proteins"

```

```

FT Region 928..935
FT /note= "proline-rich motif present in the majority
FT of streptococcal surface proteins"
FT
FT Domain 820..828
FT /label= repeat_1
FT Domain 848..857
FT /label= repeat_2
FT Domain 888..896
FT /label= repeat_3
FT Domain 927..935
FT /label= repeat_4
FT Domain 990..995
FT /note= "LPXTGX motif, found in surface proteins
FT from gram positive bacteria, with a
FT conservative replacement T to S"
XX
XX WO9506721-A.
XX
XX 09-MAR-1995.
XX
XX 30-AUG-1994; 94WO-US09926.
XX
XX 01-SEP-1993; 93US-0115227.
XX 22-AUG-1994; 94US-0115227.
XX (UVRQ ) UNIV ROCKEFELLER.
XX
XX Fischetti VA, Rakonjac J, Robbins J;
XX
XX WPI; 1995-115437/15.
XX N-PSDB; Q83181.
XX
XX A process for producing a polypeptide with apolipoproteinase
PT activity - by cloning a DNA sequence coding for a protein with
PT serum opacity factor activity and isolating said polypeptide.
XX
XX Claim 8; Fig 4A-C; 46pp; English.
XX
XX SOF22 (serum opacity factor 22) is encoded by the sof22 gene (083181)
CC from Streptococcus strain D734 and flanking regions. SOF22 has
CC apolipoproteinase activity, cleaving apolipoprotein APO-1, from high
CC density lipoprotein (HDL). Cleavage of APO-1 from HDL leads to
CC coagulation of the remaining fraction resulting in opalescence. This
CC observation is a useful tool for qualitatively determining the
CC concentration of HDL in mammalian sera, low levels of which contribute
CC to coronary heart disease.
XX
XX Sequence 1025 AA;

Query Match 27.8%; Score 49; DB 16; Length 1025;
Best Local Similarity 37.0%; Pred. No. 1.4e+02;
Matches 10; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

Qy 7 KEDSGQRDDILMEKPSRPMSNPDTPE 33
Dl 871 eedtpqrpdvllvggsdpditedtq 897

RESULT 10
ID R66407
XX R66407 standard; Protein; 563 AA.
AC R66407;
XX
XX 23-JUN-1995 (first entry)
XX
XX ELM1p protein-kinase.
XX
XX ELM1; ELM1p; protein-kinase; Elongated morphology; pseudohyphal growth;
KW pathogen transformation; polysaccharide fermentation; ethanol.
XX
XX Saccharomyces cerevisiae.
OS

```


FT Misc-difference 211 /label= unknown
FT /note= "encoded by NNN"
FT Misc-difference 212 /label= unknown
FT /note= "encoded by NNN"
FT Misc-difference 213 /label= unknown
FT /note= "encoded by NNN"
FT Misc-difference 214 /label= unknown
FT /note= "encoded by NNN"
FT Misc-difference 215 /label= unknown
FT /note= "encoded by NNN"
FT Misc-difference 216 /label= unknown
FT /note= "encoded by NAG"

PN WO9824901-A1.

XX 11-JUN-1998.

XX 25-NOV-1997; 97WO-US21606.

XX 05-DEC-1996; 96US-0760745.

XX (INCY-) INCYTE PHARM INC.

XX Bandman O, Goli SK, Murry LE;

XX WPI; 1998-333321/29.

XX N-PSDB; V39154.

XX New isolated lung growth factor variant - used to develop products
PT for treating e.g. infections, autoimmune diseases, vascular
PT conditions and cancers, or inhibiting spermatogenesis
XX
XX Claim 1; Fig 1A-C; 69pp; English.

XX This represents a human lung growth factor variant (LGFV). An expression
CC vector containing the LGFV nucleic acid can be used to transform host
CC cells for the recombinant production of the polypeptide. LGFV
CC polypeptides play a role in growth and development as well as in
CC infections, spermatogenesis, autoimmune diseases, vascular conditions and
CC cancers. LGFV polypeptides and their agonists can be used to treat cells
CC in vivo or ex vivo for tissue or organ regeneration, particularly in the
CC proliferation and differentiation of bone marrow, nerve, pancreatic or
CC renal cells. Antagonists or inhibitors of LGFV may be used to prevent
CC cancerous cell proliferation and tissue damage due to excessive
CC leukocytes produced during autoimmune or immunological responses. Such
CC cancers may include leukaemia, lymphomas or carcinomas. Excessive
CC immunological responses that may be treated include those attributed to
CC e.g. viral (AIDS), bacterial (pulmonary pneumonia, hepatitis or septic
CC shock), fungal (histoplasmosis, leprosy) or helminthic and parasitic
CC infections, allergies or asthma, arteriosclerosis, atherosclerosis or
CC collagen vascular diseases, and autoimmune diseases such as haemolytic
CC sclerosis, myasthenia gravis or rheumatoid arthritis. Inhibitors of LGFV
CC can also be used to prevent spermatogenesis in mammalian reproductive
CC tissues, thereby effecting male birth control. The products can also
CC be used for detection, diagnosis and drug screening.

XX Sequence 235 AA;

Query Match 27.8%; Score 49; DB 19; Length 235;

Best Local Similarity 38.7%; Pred. No. 24;

Matches 12; Conservative 6; Mismatches 9; Indels 4; Gaps 2;

QY 7 KEDSG---RQRDDILMEKPSRPMS-NPDTE 33

| : | | : | : | : | : | : | : |

Db 143 knekgalkrragdlledspkrpkeanpage 173

RESULT 7

R66727 R66727 standard; Protein; 240 AA.

XX AC R66727;

XX DT 02-OCT-1995 (first entry)

XX DE Human hepatoma derived growth factor.

XX KW Human hepatoma derived growth factor; recombinant production;

XX OS receptor purification.

XX PN Homo sapiens.

XX JP06343470-A.

XX PD 20-DEC-1994.

XX PF 04-JUN-1993; 93JP-0134258.

XX PR 04-JUN-1993; 93JP-0134258.

XX PA (KISH/) KISHIMOTO C.

XX PA (SEKI) SEKISUI CHEM IND CO LTD.

XX DR WPI; 1995-069304/10.

XX DR N-PSDB; Q79902, Q79903

XX PT Human hepatoma-derived growth factor DNA - also vectors and
XX PT transformed cells, useful for recombinant prodn of the growth
XX PT factor

XX PS Claim 1; Page 11; 17pp; Japanese.

XX CC Q79902 encodes R66727 human hepatoma derived growth factor. The
XX CC DNA as part of an expression vector can be used for the
XX CC recombinant production of the protein, which can be used to
XX CC encourage the growth of various animal cells, and for the
XX CC purification of receptors.

XX SQ Sequence 240 AA;

Query Match 27.8%; Score 49; DB 16; Length 240;

Best Local Similarity 38.7%; Pred. No. 24;

Matches 12; Conservative 6; Mismatches 9; Indels 4; Gaps 2;

QY 7 KEDSG---RQRDDILMEKPSRPMS-NPDTE 33

| : | | : | : | : | : | : |

Db 148 knekgalkrragdlledspkrpkeanpage 178

RESULT 8

Y69294

ID Y69294 standard; Protein; 701 AA.

XX AC Y69294;

XX DT 19-JUN-2000 (first entry)

XX DE Amino acid sequence of a Staphylococcus aureus nrdE polypeptide.

XX KW nrdE; vaccine; chromosome identification; serotyping; infection;

XX KW Helicobacter pylori; ulcer; cancer.

XX OS Staphylococcus aureus.

XX PN WO200009541-A1.

XX XX 24-FEB-2000.

XX

```

Oy 2 STRSNKEDSGRQDDILMEKPSRPMESNPDTG 34
    |||| : : | : |||| : : |||
Db 365 strskseseqnrslmek-----qdsaktdg 392

RESULT 4
W47389
ID W47389 standard; Protein; 703 AA.
XX AC
XX W47389;
XX 05-JUN-1998 (first entry)
XX DE Rice delta-1-pyrroline-5-carboxylate synthase.
XX KW Rice; delta-1-pyrroline-5-carboxylate synthase; salt resistance;
XX KW moisture loss.
XX OS Oryza sativa.
XX PN JP10057069-A.
XX PD 03-MAR-1998.
XX PF 22-AUG-1996; 96JP-0221465.
XX PR 22-AUG-1996; 96JP-0221465.
XX PA (HITA ) HITACHI LTD.
XX DR WPI; 1998-210402/19.
XX DR N-PSDB; V15744.
XX PT Rice delta-1-pyrroline-5-carboxylate synthase gene - useful for
XX PT providing salt resistance and preventing moisture loss in plants
XX PS Claim 1; Pages 3-9; 9pp; Japanese.
XX CC The present sequence is rice delta-1-pyrroline-5-carboxylate
XX CC synthase, which can be used to provide salt resistance and prevent
XX CC moisture loss in plants, improving crop yield.
XX SQ Sequence 703 AA;

Query Match 29.0%; Score 51; DB 19; Length 703;
Best Local Similarity 33.3%; Pred. No. 46;
Matches 15; Conservative 6; Mismatches 10; Indels 14; Gaps 2;

Oy 1 LSTRSNKEDSGRQ-----RDDILMEKPSRPM-----ESNPD 31
    : | : | | | | | | | | | | | | | | | | | |
Db 360 irtlanmedpinqlkkttevaddivlektscplgvlivfesrpd 404

RESULT 5
R95912
ID R95912 standard; Protein; 740 AA.
XX AC R95912;
XX DT
XX 08-NOV-1996 (first entry)
XX DE Mutant insulin-stimulated protein kinase (ISPK-1).
XX KW Insulin stimulated protein kinase; ISPK; detection; screening;
XX KW diagnosis; resistance; diabetes mellitus; NIDDM.
XX OS Homo sapiens.
XX PN WO9615233-A1.
XX PD 23-MAY-1996.
XX FT

```

```

PF 10-NOV-1995; 95WO-DK00448.
XX
PR 10-NOV-1994; 94DK-0001294.
XX
PA (NOVO ) NOVO-NORDISK AS.
XX
XX BJORBAEK C, ECHWALD SM, PEDERSEN O;
PI
XX WPI; 1996-259840/26.
DR N-PSDB; T27731.
XX
XX Mutant DNA encoding insulin-stimulated protein kinase - useful to
PT diagnose pre-disposition to insulin resistance and possibly NIDDM
XX
XX Claim 3; Page 30; 42pp; English.
XX
XX DNA encoding a mutant insulin-stimulated protein kinase (ISPK-1)
CC can be used in the detection of the presence of other mutated ISPK-1
CC genes. The method of detection comprises analysing an isolated DNA
CC by digestion with a restriction endonuclease which cleaves at the
CC site of the mutation. The obtained restriction pattern is compared
CC to a negative control restriction pattern (obtained from a wild type
CC ISPK-1 gene) and a positive control restriction pattern obtained
CC from a mutant ISPK-1 gene. This method may be used to diagnose
CC pre-disposition to insulin resistance and possibly non-insulin-
CC dependent diabetes mellitus. This sequence contains a 138S
CC substitution mutation.
XX
XX Sequence 740 AA;
SQ

Query Match 28.4%; Score 50; DB 17; Length 740;
Best Local Similarity 37.9%; Pred. No. 67;
Matches 11; Conservative 7; Mismatches 7; Indels 4; Gaps 1;

Oy 5 SNKEDSGRQDDILMEKPSRPMESNPDTG 33
    | : : | | : | : | | | | | |
Db 19 sdsaeagggq----imdepmeesnpqte 43

RESULT 6
W61963
ID W61963 standard; Protein; 235 AA.
XX
XX W61963;
XX
XX 23-SEP-1998 (first entry)
DT
XX
XX Lung growth factor variant (LGFV) protein sequence.
DE
XX
XX LGFV; lung growth factor variant; infection; autoimmune disease;
KW vascular condition; cancer; spermatogenesis inhibition; leukaemia;
KW lymphoma; carcinoma; AIDS; pulmonary pneumonia; hepatitis; asthma;
KW histoplasmosis; leprosy; parasitic infection; allergy; septic shock;
KW arteriosclerosis; atherogenesis; collagen vascular disease;
KW haemolytic sclerosis; myasthenia gravis; rheumatoid arthritis;
KW male birth control.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 207 /label= unknown
FT /note= "encoded by CNN"
FT Misc-difference 208 /label= unknown
FT /note= "encoded by NNN"
FT Misc-difference 209 /label= unknown
FT /note= "encoded by NNN"
FT Misc-difference 210 /label= unknown
FT /note= "encoded by NNN"
FT

```

PT hyper-calcaemia, hypo- or hyper-phosphatemia, kidney stones, etc
 XX
 PS Claim 9; Fig 1A-E; 62pp; English.
 XX
 CC A novel 7-transmembrane receptor (W12695) has been identified as a
 CC human G-protein parathyroid hormone (PTH) receptor, designated
 CC HTD674. It shows 48.2% homology to the human PTH receptor. Its
 CC amino acid sequence was deduced from a cDNA clone (T59619) isolated
 CC from a human T cell lymphoma tissue cDNA library. Recombinant
 CC HTD674 can be produced in transformed host cells and used to
 CC screen for (ant)agonist cpds. Agonists may be used to prevent or
 CC treat e.g. hypocalcaemia, hyperphosphataemia, hypoparathyroidism
 CC and chronic tetany by stimulating an increase in serum calcium
 CC levels. Antagonists can be used to inhibit the receptor e.g. for
 CC the treatment of osteoporosis, hypercalcaemia, hypoparathyroidism,
 CC hypophosphataemia, kidney stone, nephrolithiasis.
 XX
 SQ Sequence 541 AA;

Query Match 100.0%; Score 176; DB 18; Length 541;
 Best Local Similarity 100.0%; Pred. No. 1.2e-16;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSTRNKEDSGRQDDILMEKPSRPMESNPDETEG 34
 |||||
 DB 508 lstrnkedsgrqddilmebpsrmpesnpdteg 541

RESULT 2
 W48720
 ID W48720 standard; Protein; 993 AA.
 XX
 AC W48720;
 XX
 DT 19-AUG-1998 (first entry)
 XX
 DE Human mitochondrial isoleucyl-tRNA synthetase.
 XX
 KW Mitochondrial isoleucyl-tRNA synthetase; T-cell; pathogen.
 XX
 OS Homo sapiens.
 XX
 PN US5759833-A.
 XX
 XX 02-JUN-1998.
 XX
 XX 27-MAY-1994; 94US-0250852.
 XX
 PR 06-JUN-1995; 95US-0468557.
 PR 27-MAY-1994; 94US-0250852.
 XX
 PA (CANC-) CANCER INST JAPANESE FOUND CANCER.
 PA (CUBI-) CUBIST PHARM INC.
 XX
 PI Kranz JE, Schimmel PR, Shiba K;
 XX
 DR WPI; 1998-332142/29.
 DR N-PSDB; V18326.
 XX
 XX Human isoleucyl-tRNA synthetase DNA - useful for producing
 PT recombinant proteins, tester strains, etc.
 XX
 PS Claim 3; Columns 47-52; 46pp; English.
 XX
 CC The present sequence is that of a human mitochondrial isoleucyl-tRNA
 CC synthetase. The cDNA encoding for the mitochondrial isoleucyl-tRNA
 CC synthetase was isolated from a human T-cell cDNA library. The cDNA
 CC can be expressed using expression vectors to produce the corresponding
 CC recombinant protein. The protein can be used in assays to test
 CC substances known to inhibit the isoleucyl-tRNA synthetase or other tRNA
 CC synthetases of pathogenic organisms. Also, expression of the
 CC isoleucyl-tRNA synthetase in a tester strain can be useful for testing

CC substances which are capable of inhibiting its activity.
 XX
 SQ Sequence 993 AA;

Query Match 34.1%; Score 60; DB 19; Length 993;
 Best Local Similarity 38.9%; Pred. No. 3.8;
 Matches 14; Conservative 5; Mismatches 13; Indels 4; Gaps 1;

QY 2 STRSNKEDSGRQDDILMEKPSRPMESNPDETEG 33
 |||||
 DB 30 shqpsnsgrqddilvlpqtsfpmkilgrqqpde 65

RESULT 3
 Y68822
 ID Y68822 standard; protein; 1589 AA.
 XX
 AC Y68822;
 XX

DT 16-MAY-2000 (first entry)

DE Amino acid sequence of a Son of sevenless (Sos) related protein.

KW Ras; Son of sevenless; Sos; crystal; Ras-Sos complex; cancer;

KW protein coordinate data.

OS Saccharomyces cerevisiae.

PN W0200005258-A1.

XX 03-FEB-2000.

XX 20-JUL-1999; 99WO-US16348.

XX 21-JUL-1998; 98US-0119794.

XX (UYRQ) UNIV ROCKEFELLER.

XX Boriack-Sjodin A, Margarit SM, Bar-Sagi D, Cole P, Kuriyan J;

XX WPI; 2000-182647/16.

XX Novel crystals comprising a Ras-Son of sevenless complex, useful for
 PT screening drugs useful in cancer treatment

XX Disclosure; Page 175-181; 224pp; English.

XX The specification describes a crystal complex comprising at least a
 CC Sos contacting region of a Ras protein and at least a Son of sevenless
 CC (Sos) protein catalytic region fragment, that effectively diffracts
 CC X-rays. Ras and Sos form a tight complex. Sos does not impede the
 CC binding sites for the nucleotide base and the ribose of GTP or GDP
 CC and thus the Ras-Sos complex maintains a structure that permits
 CC nucleotide release and rebinding. The crystals are used for the
 CC determination of the atomic coordinates of the complex to a resolution
 CC of more than 5.0 Angstrom. The crystals, or a dataset comprising the
 CC three-dimensional coordinates obtained from the crystals, is useful
 CC for identifying an agent that stabilizes the Ras-Sos complex. The
 CC crystals are also useful for identifying agents that inhibit the
 CC formation of Ras-Sos complex. Ras and Sos fragments are useful for
 CC growing a crystal of a protein-ligand complex. Agents that stabilize
 CC or inhibit the formation of Ras-Sos complex are useful in the
 CC treatment of cancer. The present sequence represents a yeast
 CC Sos-related protein.

XX Sequence 1589 AA;

Query Match 31.0%; Score 54.5; DB 21; Length 1589;
 Best Local Similarity 36.4%; Pred. No. 39;
 Matches 12; Conservative 8; Mismatches 8; Indels 5; Gaps 1;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:51:52 ; Search time 138.73 Seconds
(without alignments)
8.380 Million cell updates/sec

Title: US-09-236-468A-2_COPY_508_541

Perfect score: 176

Sequence: 1 LSTRSNKESGRQDILMEKPSRPMESNPDTG 34

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_36.*

- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	176	100.0	541	18 W12695	G-protein parathyr
2	60	34.1	193	19 W48720	Human mitochondria
3	54.5	31.0	1589	21 Y68822	Amino acid sequenc
4	51	29.0	703	19 W47389	Rice delta-i-pyrro
5	50	28.4	740	17 R95912	Mutant insulin-sti
6	49	27.8	235	19 W61963	Lung growth factor
7	49	27.8	240	16 R66727	Human hepatoma der
8	49	27.8	701	21 Y69294	Amino acid sequenc
9	49	27.8	1025	16 R70126	Serum opacity fact
10	48.5	27.6	563	16 R66407	ELMlp protein-kin
11	48	27.3	730	20 Y13459	Amino acid sequenc
12	47.5	27.0	237	18 W33216	Mouse hepatoma der

13	47	26.7	223	20 W98181	Alfalfa cyclin-dep
14	47	26.7	456	20 Y60239	Human endometrium
15	46	26.1	321	13 R27118	Fucose dehydrogena
16	46	26.1	826	13 R26042	P. yoelii SSP2 ant
17	46	26.1	1041	20 Y41768	Human PRO286 prote
18	46	26.1	1041	20 Y05867	Human toll protein
19	45.5	25.9	881	20 Y11681	Sulfated fucose-co
20	45	25.6	170	19 W61546	Novel secreted pro
21	45	25.6	208	20 W93590	Human TNRL3 protei
22	45	25.6	231	19 W49876	Human polyhomeotic
23	45	25.6	231	19 W59426	Homo sapiens polyh
24	45	25.6	249	19 W29745	TNF related endoth
25	45	25.6	249	20 Y09369	Human tumour necro
26	45	25.6	284	19 W47525	Homo sapiens tumou
27	45	25.6	312	20 Y07109	Colon cancer assoc
28	45	25.6	352	21 Y32206	Human receptor mol
29	45	25.6	444	20 W90340	G. max truncated S
30	45	25.6	516	7 P61362	Soybean glycinin A
31	45	25.6	524	20 W90339	G. max SBP1 protei
32	45	25.6	1093	14 R42818	TMF. Homo sapiens
33	44.5	25.3	603	19 W83214	Human h-NUMB. Hom
34	44	25.0	38	15 R46631	Endoproteinase Asp
35	44	25.0	282	21 Y58635	Protein regulating
36	44	25.0	315	20 Y34511	Porphyrinomonas ging
37	44	25.0	329	20 Y34384	Porphyrinomonas ging
38	44	25.0	439	19 W59995	Cell division Div1
39	44	25.0	451	19 W69951	RANK polypeptide p
40	44	25.0	451	19 W68287	RANK polypeptide p
41	44	25.0	509	19 W40480	Rat SH2 binding pr
42	44	25.0	532	20 Y16036	Neurospora crassa
43	44	25.0	591	19 W69950	RANK polypeptide p
44	44	25.0	591	19 W68286	RANK polypeptide p
45	44	25.0	615	20 W30659	Human tumour necro

ALIGNMENTS

RESULT 1

W12695

ID W12695 standard: Protein; 541 AA.

XX W12695;

XX 31-MAY-1997 (first entry)

XX G-protein parathyroid hormone receptor HLTG74.

XX G-protein parathyroid hormone receptor; HLTG74; parathormone; PTH;
KW calcium; signal transduction; agonist; antagonist; hypocalcaemia;
KW hyperphosphataemia; hypoparathyroidism; chronic tetany;
KW osteoporosis; hypercalcaemia; hypoparathyroidism; hypophosphataemia;
KW kidney stone; nephrolithiasis; therapy; diagnosis.

XX Homo sapiens.

XX WO9639433-A1.

XX 12-DEC-1996.

XX 05-JUN-1995; 95WO-US07085.

XX 05-JUN-1995; 95WO-US07085.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Li Y, Rosen CA, Ruben SM, Soppet DR;

XX WPI: 1997-043068/04.

XX N-PSDB; T59619.

XX Human G-protein parathyroid hormone receptor, HLTG74 - used to

XX identify (ant)agonists, used in the treatment of hypo- or

XX PT

DR INTERPRO; IPR000169; -.
DR INTERPRO; IPR000668; -.
DR PFAM; PF00112; Peptidase_C1; 1.
DR PRINTS; PR00705; PAPA1N.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
DR PROSITE; PS00640; THIOL_PROTEASE_ASN; 1.
KW Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 360 PRE-PRO-TPE4A PROTEIN.
SQ SEQUENCE 360 AA; 40685 MW; 55A1B49E4108213C CRC64;

Query Match 71.4%; Score 30; DB 10; Length 360;
Best Local Similarity 85.7%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVTQSR 8
 |||||
Db 295 GVTQDRT 301

Search completed: November 8, 2000, 08:56:22
Job time: 533 sec

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarly C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003732; AAF55803.1; -;
 DR FLYBASE; FBgn0038824; CG15689.
 SQ SEQUENCE 195 AA; 21409 MW; 700FD33811CD03FC CRC64;

Query Match 71.4%; Score 30; DB 5; Length 195;
 Best Local Similarity 71.4%; Pred. No. 39;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVTQSR 8
 I:|||||
 Db 3 GVTESRT 9

RESULT 13
 O14596 PRELIMINARY; PRT; 245 AA.
 ID O14596;
 AC O14596;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)
 DE FOLATE BINDING PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RC TISSUE-SALIVARY GLAND;
 RA Verna R.S., Elwood P.C.;
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF000380; AAB81937.1; -;
 SQ SEQUENCE 245 AA; 26268 MW; BA36B7DB95FC0E20 CRC64;

Query Match 71.4%; Score 30; DB 4; Length 245;
 Best Local Similarity 66.7%; Pred. No. 49;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGVTOSRTA 9
 |||||:|
 Db 203 SGVTHRSA 211

RESULT 14
 Q41697 PRELIMINARY; PRT; 359 AA.
 ID Q41697;
 AC Q41697;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE CYSTEINE PROTEINASE PRECURSOR.
 OS *Vicia sativa* (Spring vetch) (Tare).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Fabales; Fabaceae;
 OC Papilionoideae; *Vicia*.

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=COTYLEDON;
 RX MEDLINE; 98004268.
 RA Becker C., Senyuk V.J., Shutov A.D., Nong V.H., Fischer J.,
 RA Horstmann C., Muentz K.;
 RA "Proteinase A, a storage-globulin-degrading endopeptidase of vetch-
 RT (*Vicia sativa* L.) seeds, is not involved in early steps of storage-
 RT protein mobilization.";
 RL Eur. J. Biochem. 248:304-312(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=COTYLEDON;
 RX MEDLINE; 78061037.
 RA Bul'maga V., Shutov A.;
 RT "[Partial purification and characterization of protease A of
 RT germinating vetch seeds, hydrolyzing native reserve proteins].";
 RL Biochimia 42:1983-1989(1977).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=COTYLEDON;
 RA Shutov A., Vaintraub I.;
 RL Phytochemistry 26:1557-1566(1987).
 DR EMBL; Z34895; CAA84378.1; -;
 DR HSSP; P07711; 1CJL.
 DR MENDEL; 12437; Vicsa; 1134; 12437.
 DR INTERPRO; IPR000169; -;
 DR INTERPRO; IPR000668; -;
 DR INTERPRO; IPR000886; -;
 DR PFAM; PF00112; Peptidase_C1; 1.
 DR PRINTS; PR00705; PAPAIN.
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
 DR PROSITE; PS00640; THIOL_PROTEASE_ASN; 1.
 KW Signal.
 FT SIGNAL. 1 20 POTENTIAL.
 FT CHAIN 128 359 CYSTEINE PROTEINASE.
 SQ SEQUENCE 359 AA; 40632 MW; C774975F51292592 CRC64;

Query Match 71.4%; Score 30; DB 10; Length 359;
 Best Local Similarity 85.7%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVTQSR 8
 |||||
 Db 294 GVTQDRT 300

RESULT 15
 O82708 PRELIMINARY; PRT; 360 AA.
 ID O82708;
 AC O82708;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE PRE-PRO-TPE4A PROTEIN PRECURSOR.
 GN TPE4A.
 OS *Pisum sativum* (Garden pea).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Fabales; Fabaceae;
 OC Papilionoideae; *Pisum*.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. ALASKA;
 RA Cercos M., Santamaria S., Carbonell J.;
 RT "TPE4A: A thiol-protease gene induced during ovary senescence and seed
 RT germination in *Pisum sativum*.";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ004958; CAA06243.1; -;
 DR HSSP; P07711; 1CJL.
 DR MENDEL; 33280; Pissa; 1134; 33280.

RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z68005; CA91994.1; -;
DR EMBL; Z68006; CA91994.1; JOINED.
DR EMBL; Z68006; CA91999.1; -;
DR EMBL; Z68005; CA91999.1; JOINED.
DR HSSP; P05164; 1MHL.
DR INTERPRO; IPR000372; -;
DR INTERPRO; IPR001536; -;
DR INTERPRO; IPR001611; -;
DR INTERPRO; IPR002007; -;
DR INTERPRO; IPR002016; -;
DR INTERPRO; IPR003006; -;
DR PFAM; PF00047; 1g; 2;
DR PFAM; PF00141; peroxidase; 2.
DR PFAM; PF00560; LRR; 4.
DR PFAM; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR PRINTS; PR00457; ANPEROXIDASE.
SQ SEQUENCE 1328 AA; 151150 MW; E0812BE92C81723A CRC64;

Query Match 73.8%; Score 31; DB 5; Length 1328;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVTSRTA 9

Db 1151 GVTEKRTA 1158

RESULT 10

Q9PYF2 ID Q9PYF2 PRELIMINARY; PRT; 2185 AA.

AC Q9PYF2;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)

DE POLYPROTEIN.

OS Cossackievirus B5.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Enterovirus.

[1]

RN SEQUENCE FROM N.A.

RC STRAIN=FAULKNER;

RA Lindberg A.M., Polacek C.;

RT "Molecular analysis of the prototype cossackievirus B5 genome.";

RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF114383; AAF21971.1; -;

DR INTERPRO; IPR000081; -;

DR INTERPRO; IPR000199; -;

DR INTERPRO; IPR000605; -;

DR INTERPRO; IPR001205; -;

DR INTERPRO; IPR001643; -;

DR INTERPRO; IPR001676; -;

DR PFAM; PF00073; rhv; 3.

DR PFAM; PF00548; Cys-protease-3C; 1.

DR PFAM; PF00680; RNA_dep_RNA_pol; 1.

DR PFAM; PF00910; RNA_helicase; 1.

DR PFAM; PF00947; PICO_P2A; 1.

DR PRINTS; PR00918; CALICIVIRUSNS.

KW Polyprotein.

SQ SEQUENCE 2185 AA; 243318 MW; BFA325AEDC17567D CRC64;

Query Match 73.8%; Score 31; DB 12; Length 2185;

Best Local Similarity 75.0%; Pred. No. 2.6e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SGVTSRT 8

Db 836 SGVTEGRT 843

RESULT 11

Q9WRW8 ID Q9WRW8 PRELIMINARY; PRT; 130 AA.

AC Q9WRW8;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)

DE VP1 (FRAGMENT).

OS Echovirus 33.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Enterovirus.

[1]

RN SEQUENCE FROM N.A.

RC STRAIN=PER98-2558;

RX MEDLINE; 99221739.

RA Oberste M.S., Maher K., Kilpatrick D.R., Flemister M.R., Brown B.A.,

RA Pallansch M.A.;

RT "Typing of human enteroviruses by partial sequencing of VP1.";

RL J. Clin. Microbiol. 37:1288-1293(1999).

DR EMBL; AF081632; AAD40536.1; -;

DR INTERPRO; IPR001676; -;

DR PFAM; PF00073; rhv; 1.

DR NON_TER 1

FT NON_TER 130 130

SQ SEQUENCE 130 AA; 14607 MW; CF490C3F23C3F27B CRC64;

Query Match 71.4%; Score 30; DB 12; Length 130;

Best Local Similarity 85.7%; Pred. No. 26;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVTSRT 8

Db 119 GVTSRT 125

RESULT 12

Q9VDJ0 ID Q9VDJ0 PRELIMINARY; PRT; 195 AA.

AC Q9VDJ0;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)

DE CGI5689 PROTEIN.

GN CGI5689.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

[1]

RN SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RX MEDLINE; 20196006.

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Fostier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

```

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE CAPSID PROTEIN (FRAGMENT).
GN VP1.
OS Echovirus 13.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DEL CARMEN;
RA Bailey J.L., Bequet A., Chambon M., Henquell C., Peigue-Lafeuille H.;
RT "A specific RT-PCR using anchored primers for human enteroviruses
RT Species B : application to the molecular epidemiology of echoviruses
RT and to the study of nosocomial infection with echovirus type 30 in a
RT neonatal unit.";
RL Infected. Immun. 66:3810-3817(1998).
DR EMBL; AJ241427; CAB53937.1; -.
DR INTERPRO; IPR001676; -.
DR PFAM; PF00073; rhv; 1.
FT NON_TER 287 287
FT NON_TER 287 287
SQ SEQUENCE 287 AA; 31684 MW; 32A57A7D71F72CEA CRC64;

Query Match 76.2%; Score 32; DB 12; Length 287;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVTQSRTA 9
Db 265 GVTESRTS 272

RESULT 6
ID O69413 PRELIMINARY; PRT; 374 AA.
AC O69413;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DE 01-AUG-1998 (TREMBlrel. 07, Last annotation update)
DE ESPD PROTEIN.
GN ESPD.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EHEC EDL933;
RA Kresse A.U., Ebel F., Deibel C., Chakraborty T., Guzman C.A.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y13068; CAA73507.1; -.
SQ SEQUENCE 374 AA; 39093 MW; F73B6B950C5FF3D4 CRC64;

Query Match 76.2%; Score 32; DB 2; Length 374;
Best Local Similarity 75.0%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGVTQSRT 8
Db 22 SGITQSET 29

RESULT 7
ID O85626 PRELIMINARY; PRT; 374 AA.
AC O85626;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE L0021.
GN ESPD.
OS Escherichia coli.

```

```

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EDL933;
RX MEDLINE; 98333985.
RA Perna N.T., Mayhew G.F., Posfai G., Elliott S., Donnenberg M.S.,
RA Kaper J.B., Blattner F.R.;
RT "Molecular evolution of a pathogenicity island from enterohemorrhagic
RT Escherichia coli O157:H7.";
RL Infect. Immun. 66:3810-3817(1998).
DR EMBL; AF071034; AAC31500.1; -.
SQ SEQUENCE 374 AA; 39083 MW; F739BB94DC5CF3D4 CRC64;

Query Match 76.2%; Score 32; DB 2; Length 374;
Best Local Similarity 75.0%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGVTQSRT 8
Db 22 SGITQSET 29

RESULT 8
ID P74816 PRELIMINARY; PRT; 670 AA.
AC P74816;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE HYPOTHETICAL 69.5 KDA PROTEIN.
GN SPSR.
OS Sphingomonas sp. S88.
OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonas group;
OC Sphingomonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S88;
RX MEDLINE; 96196177.
RA Yamazaki M., Thorne L., Mikolajczyk M., Armentrout R.W., Pollock T.J.;
RT "Linkage of genes essential for synthesis of a polysaccharide capsule
RT in Sphingomonas strain S88.";
RL J. Bacteriol. 178:2676-2687(1996).
DR EMBL; U51197; AAC44057.1; -.
KW Hypothetical protein.
SQ SEQUENCE 670 AA; 69534 MW; 768E04D60B9705E4 CRC64;

Query Match 73.8%; Score 31; DB 2; Length 670;
Best Local Similarity 66.7%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGVTQSRTA 9
Db 307 SGITQSNA 315

RESULT 9
ID Q21043 PRELIMINARY; PRT; 1328 AA.
AC Q21043; Q21390;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE K09C8.5 PROTEIN.
GN K09C8.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peleoderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA Kershaw J.;

```


Q9VFM8
ID Q9VFM8 PRELIMINARY; PRT; 261 AA.
AC Q9VFM8
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE 140UP PROTEIN.
GN 140UP.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE; 20196006.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananatzides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Sutton G.G., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Fodor B., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heilmann T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Relnert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL; AF003703; AAF55023.1; -.
DR FLYBASE; FBgn0010340; 140up.
SQ SEQUENCE 261 AA; 29182 MW; 5DB78CF6CFC4435A CRC64;

Query Match 78.6%; Score 33; DB 5; Length 261;
Best Local Similarity 87.5%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVTSORTA 9
DB 86 GVTSORVA 93
|||:|||||

RESULT 3
ID O85742 PRELIMINARY; PRT; 172 AA.
AC O85742;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE RSBU (FRAGMENT).
GN RSBU.
OS Listeria monocytogenes
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Listeria.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LO4035;
RX MEDLINE; 98389670.
RA Becker L.A., Cetin M.S., Hutkins R.W., Benson A.K.;
RT "Identification of the gene encoding the alternative sigma factor
RT sigmaB from Listeria monocytogenes and its role in osmotolerance.";
RL J. Bacteriol. 180:4547-4554(1998).
DR EMBL; AF074855; AAC34824.1; -.
FT NON_TER 1
SQ SEQUENCE 172 AA; 19762 MW; 2C775E008EB44DEF CRC64;

Query Match 76.2%; Score 32; DB 2; Length 172;
Best Local Similarity 75.0%; Pred. No. 12;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVTSORTA 9
DB 112 GVTSORTA 119
|||:|||||

RESULT 4
ID Q9YLL4 PRELIMINARY; PRT; 287 AA.
AC Q9YLL4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CAPSID PROTEIN VP1 (FRAGMENT).
GN VP1.
OS Echovirus 13.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DEL CARMEN;
RX MEDLINE; 99138973.
RA Oberste M.S., Maher K., Kilpatrick D.R., Pallansch M.A.;
RT "Molecular evolution of the human enteroviruses: correlation of
RT serotype with VP1 sequence and application to picornavirus
RT classification";
RL J. Virol. 73:1941-1948(1999).
DR EMBL; AF081327; AAD17731.1; -.
DR HSP; P03313; ICOV.
DR INTERPRO; IPR001676; -.
DR PFAM; PF000073; rhv; 1.
FT NON_TER 1
FT NON_TER 287
SQ SEQUENCE 287 AA; 31783 MW; AD1A5B540FE2A8E0 CRC64;

Query Match 76.2%; Score 32; DB 12; Length 287;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVTSORTA 9
DB 265 GVTSORTS 272
|||:|||||

RESULT 5
ID Q9QP54 PRELIMINARY; PRT; 287 AA.
AC Q9QP54;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:56:20 ; Search time 152.43 Seconds
(without alignments)
5.513 Million cell updates/sec

Title: US-09-236-468a-2_COPY_496_504
Perfect score: 42
Sequence: 1 SGVTQSRTA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 297973

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_14:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	78.6	261	5 P81928	P81928 drosophila
2	33	78.6	261	5 Q9VEM8	Q9VEM8 drosophila
3	32	76.2	172	2 O85742	O85742 listeria mo
4	32	76.2	287	12 Q9YL14	Q9YL14 echovirus 1
5	32	76.2	287	12 Q9QP54	Q9QP54 echovirus 1
6	32	76.2	374	2 O69413	O69413 escherichia
7	32	76.2	374	2 O85626	O85626 escherichia
8	31	73.8	670	2 P74816	P74816 sphingomona
9	31	73.8	1328	5 Q21043	Q21043 caenorhabdi
10	31	73.8	2185	12 Q9PYF2	Q9PYF2 coxsackievi
11	30	71.4	130	12 Q9WRW8	Q9WRW8 echovirus 3
12	30	71.4	195	5 Q9VDJ0	Q9VDJ0 drosophila
13	30	71.4	245	4 O14596	O14596 homo sapien
14	30	71.4	359	10 O41697	O41697 vicia sativ
15	30	71.4	360	10 O82708	O82708 pisum sativ
16	30	71.4	366	5 Q9V704	Q9V704 drosophila
17	30	71.4	445	2 O83891	O83891 treponema p
18	30	71.4	515	12 Q64784	Q64784 avian adeno
19	30	71.4	866	5 Q9VCV5	Q9VCV5 drosophila

20	30	71.4	968	5 Q9VCB2	Q9VCB2 drosophila
21	30	71.4	1021	5 O15733	O15733 dictyosteli
22	30	71.4	1061	9 O64335	O64335 bacterioph
23	29	69.0	123	10 Q9SXG0	Q9SXG0 oryza sativ
24	29	69.0	133	12 Q9W9F8	Q9W9F8 coxsackievi
25	29	69.0	138	12 Q9WAC6	Q9WAC6 echovirus 6
26	29	69.0	177	11 Q9WUC0	Q9WUC0 mus musculu
27	29	69.0	231	2 Q9XCK3	Q9XCK3 streptococc
28	29	69.0	248	12 O89551	O89551 sugarcane s
29	29	69.0	263	2 O52313	O52313 streptomyce
30	29	69.0	289	12 Q9YID5	Q9YID5 echovirus 6
31	29	69.0	289	12 Q9YID4	Q9YID4 echovirus 6
32	29	69.0	289	12 Q9YID3	Q9YID3 echovirus 6
33	29	69.0	303	13 Q9PTI7	Q9PTI7 gallus gall
34	29	69.0	421	1 O28350	O28350 archaeeoglob
35	29	69.0	462	4 O60717	O60717 homo sapien
36	29	69.0	475	4 O60492	O60492 homo sapien
37	29	69.0	475	4 Q9UBM7	Q9UBM7 homo sapien
38	29	69.0	547	5 Q20953	Q20953 caenorhabdi
39	29	69.0	815	5 O44385	O44385 caenorhabdi
40	29	69.0	1110	5 O09493	O09493 caenorhabdi
41	29	69.0	1440	5 O27395	O27395 caenorhabdi
42	28	66.7	49	12 Q9WAM8	Q9WAM8 hepatitis c
43	28	66.7	88	2 P97023	P97023 brevibacter
44	28	66.7	95	12 Q9QNT7	Q9QNT7 human immun
45	28	66.7	99	2 O85015	O85015 listeria mo

ALIGNMENTS

RESULT 1

P81928

ID P81928 PRELIMINARY; PRT; 261 AA.

AC P81928;

DT 01-AUG-1999 (TREMBLrel. 11, Created)

DT 01-AUG-1999 (TREMBLrel. 11, Last sequence update)

DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)

DE RP1140-UPSTREAM PROTEIN.

GN 1400P OR RP1140 UPSTREAM.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=EMBRYO;

RX MEDLINE; 91276237.

RA Sitzler S., Oldenburg I., Petersen G., Bautz E.K.F.;

RT "Analysis of the promoter region of the housekeeping gene DmRP140 by sequence comparison of Drosophila melanogaster and Drosophila virilis."

RL Gene 100:155-162(1991).

CC -1- FUNCTION: ESSENTIAL FOR VIABILITY.

DR EMBL; M62975; AAD40352.2; -.

KW FLYBASE; FBgn0010340; 140up.

FT TRANSMEM 131 151

FT TRANSMEM 183 203

FT CARBOHYD 31 31

SQ SEQUENCE 261 AA; 29242 MW; 060D2C47DFC92E26 CRC64;

Query Match 78.6%; Score 33; DB 5; Length 261;
Best Local Similarity 87.5%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVTQSRTA 9
| | | | | | |
DB 86 GVTQSRVA 93

RESULT 2

Qy 2 GVTQSR 7
|||||
Db 194 GVTQSR 199

```
RESULT 15
CENB_MOUSE
ID CENB_MOUSE STANDARD; PRT; 599 AA.
AC P27790;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE MAJOR CENTROMERE AUTOANTIGEN B (CENTROMERE PROTEIN B) (CENP-B).
GN CENPB OR CENP-B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=LIVER;
RX MEDLINE; 91372020.
RA Sullivan K.F., Glass C.A.;
RT "CENP-B is a highly conserved mammalian centromere protein with
RT homology to the helix-loop-helix family of proteins.";
RL Chromosoma 100:360-370(1991).
CC -!- FUNCTION: INTERACTS WITH CENTROMERIC HETEROCHROMATIN IN
CC CHROMOSOMES AND BINDS TO A SPECIFIC SUBSET OF ALPHOID SATELLITE
CC DNA, CALLED THE CENP-B BOX. MAY ORGANIZE ARRAYS OF CENTROMERE
CC SATELLITE DNA INTO A HIGHER ORDER STRUCTURE WHICH THEN DIRECTS
CC CENTROMERE FORMATION AND KINETOCORE ASSEMBLY IN MAMMALIAN
CC CHROMOSOMES (BY SIMILARITY).
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: STRONG, WITH OTHER MAMMALIAN CENP-B.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X5038; CAA38878.1; -.
CC MGD; MGI:88376; CENPB.
CC Chromosomal protein; Nuclear protein; DNA-binding; Centromere.
CC DNA_BIND 1 125
CC FT DOMAIN 404 465 GLU-RICH (ACIDIC).
CC FT DOMAIN 508 538 ASP/GLU-RICH (ACIDIC).
CC SQ SEQUENCE 599 AA; 65478 MW; 07C1CCE6CC5B2185 CRC64;
```

Query Match 69.08; Score 29; DB 1; Length 599;
Best Local Similarity 85.7%; Pred. No. 65;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SGVTQSR 7
|||||
Db 136 SGVTRSR 142

Search completed: November 8, 2000, 09:03:58
Job time: 868 sec

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; 297369; CAB10596.1; -.
DR HSSP; P03617; 1IFD.
DR INTERPRO; IPR000897; -.
DR PFAM; PF00448; SRP54; 1.
DR PROSITE; PS00300; SRP54; 1.
KW Signal recognition particle; GTP-binding; RNA-binding; Membrane;
KW Cell division.

FT NP_BIND 238 245 GTP (BY SIMILARITY).
FT NP_BIND 320 324 GTP (BY SIMILARITY).
FT NP_BIND 382 385 GTP (BY SIMILARITY).
SQ SEQUENCE 430 AA; 45354 MW; E9F78589AAB2C598 CRC64;

Query Match 69.0%; Score 29; DB 1; Length 430;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGVTQSTRA 9
Db 57 SGITFSQTA 65

RESULT 13

HVCS_HETFR
ID HVCS_HETFR STANDARD; PRT; 438 AA.
AC P23087;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN C REGION, SECRETED FORM (CLONE 3050).
OS Heterodontus francisci (Horn shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Heterodontoidae; Heterodontus.
OC Heterodontus.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-SPLEEN;
RA Kokubu F., Hinds K., Litman R., Shambloott M.J., Litman G.W.;
RT "Complete structure and organization of immunoglobulin heavy chain
constant region genes in a phylogenetically primitive vertebrate.";
RL EMBO J. 7:1979-1988(1988).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X07781; CAA30613.1; ALT_INIT.

DR PIR; S01853; HVRKCS.
DR INTERPRO; IPR000495; -.
DR PFAM; PF00047; Ig. 4.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;

FT NON_TER 1 99
FT DOMAIN 1 99 CH1.
FT DOMAIN 100 205 CH2.
FT DOMAIN 206 308 CH3.
FT DOMAIN 309 418 CH4.
FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 411 411 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 415 415 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 438 AA; 47920 MW; 9BDF37C3BBD7D15B CRC64;

Query Match 69.0%; Score 29; DB 1; Length 438;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVTQSR 7
Db 194 GVTQSR 199

RESULT 14

HVCM_HETFR
ID HVCM_HETFR STANDARD; PRT; 461 AA.
AC P23088;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN C REGION, MEMBRANE-BOUND FORM (CLONE 3050).
OS Heterodontus francisci (Horn shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Heterodontoidae; Heterodontidae;
OC Heterodontus.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-SPLEEN;
RA Kokubu F., Hinds K., Litman R., Shambloott M.J., Litman G.W.;
RT "Complete structure and organization of immunoglobulin heavy chain
constant region genes in a phylogenetically primitive vertebrate.";
RL EMBO J. 7:1979-1988(1988).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X07781; CAA30614.1; ALT_SEQ.

DR PIR; S01854; HVRKCO.
DR INTERPRO; IPR000495; -.
DR PFAM; PF00047; Ig. 4.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;

FT NON_TER 1 99
FT DOMAIN 1 99 CH1.
FT DOMAIN 100 205 CH2.
FT DOMAIN 206 308 CH3.
FT DOMAIN 309 418 CH4.
FT TRANSMEM 438 458 POTENTIAL.

FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 415 415 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 461 AA; 50762 MW; 029E91C88D5DF911 CRC64;

Query Match 69.0%; Score 29; DB 1; Length 461;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
SQ SEQUENCE 307 AA; 30126 MW; 9C7206CC18660F0B CRC64;

Query Match 69.0%; Score 29; DB 1; Length 307;
Best Local Similarity 85.7%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGVTQSR 7
DB 73 SGVTRSR 79

RESULT 10
Y412_MYCPN STANDARD; PRT; 372 AA.
AC P75184;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL LIPOPROTEIN MG412 HOMOLOG PRECURSOR.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE; 97105885.
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
CC ANCHOR (POTENTIAL).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AE000023; AAB95879.1; ALT_INIT.
CC INTERPRO; IPR002370; -.
CC PFAM; PF01449; PstS; 1.
CC PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
CC KW Hypothetical protein; Lipoprotein; Signal.
CC FT SIGNAL 1 22 POTENTIAL.
CC FT CHAIN 23 372 HYPOTHETICAL LIPOPROTEIN MG412 HOMOLOG.
CC FT LIPID 23 23 N-ACYL DIGLYCERIDE (POTENTIAL).
CC SQ SEQUENCE 372 AA; 40976 MW; 608D661025BA662B CRC64;

Query Match 69.0%; Score 29; DB 1; Length 372;
Best Local Similarity 62.5%; Pred. No. 39;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGVTQSR 8
DB 190 SGLTQTKT 197

RESULT 11
HVC3_HETFR STANDARD; PRT; 393 AA.
AC P23086;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN C REGION (CLONE 6121) (FRAGMENT).
OS Heterodontus francisci (Horn shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
```

```
OC Elasmobranchii; Galeomorphii; Heterodontoidea; Heterodontidae;
OC Heterodontus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RX MEDLINE; 88328985.
RA Kokubu F., Hinds K., Litman R., Shambloott M.J., Litman G.W.;
RT "Complete structure and organization of immunoglobulin heavy chain
RL EMBO J. 7:1979-1988(1988).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X07782; CAA30615.1; -.
CC PIR; S01852; HVRKCI.
CC INTERPRO; IPR000495; -.
CC INTERPRO; IPR003006; -.
CC PFAM; PF00047; Ig; 3.
CC DR PROSITE; PS00290; IG_MHC; 3.
CC KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
CC FT NON_TER 1 1
CC FT CARBOHYD 119 119 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 366 366 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 370 370 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 380 380 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SQ SEQUENCE 393 AA; 43081 MW; 4E4D076972F18B5 CRC64;

Query Match 69.0%; Score 29; DB 1; Length 393;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVTQSR 7
DB 149 GVTQSR 154

RESULT 12
FTSY_MYCLE STANDARD; PRT; 430 AA.
AC O33010;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CELL DIVISION PROTEIN FTSY HOMOLOG.
GN FTSY OR MLCB250.02.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RA Seeger K.J., Harris D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: FUNCTIONAL HOMOLOG OF SRP RECEPTOR. PROBABLY INVOLVED IN
CC THE RECEPTION AND INSERTION OF A SUBSET OF PROTEINS AT THE
CC MEMBRANE (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SRP FAMILY OF GTP-BINDING PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
DR  WORMPEP; W08D2.5; CE06560.
DR  INTERPRO; IPR001757; -.
DR  PFAM; PF00122; E1-E2_ATPase; 4.
DR  PROSITE; PS00154; ATPASE_E1_E2; 1.
KW  Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding.
FT  TRANSMEM 1 15 POTENTIAL.
FT  DOMAIN 16 21 EXTRACELLULAR (POTENTIAL).
FT  TRANSMEM 22 44 POTENTIAL.
FT  DOMAIN 45 181 CYTOPLASMIC (POTENTIAL).
FT  TRANSMEM 182 204 POTENTIAL.
FT  DOMAIN 205 207 EXTRACELLULAR (POTENTIAL).
FT  TRANSMEM 208 226 POTENTIAL.
FT  DOMAIN 227 377 CYTOPLASMIC (POTENTIAL).
FT  TRANSMEM 378 397 POTENTIAL.
FT  DOMAIN 398 410 EXTRACELLULAR (POTENTIAL).
FT  TRANSMEM 411 432 POTENTIAL.
FT  DOMAIN 433 937 CYTOPLASMIC (POTENTIAL).
FT  TRANSMEM 938 957 POTENTIAL.
FT  DOMAIN 958 964 EXTRACELLULAR (POTENTIAL).
FT  TRANSMEM 965 982 POTENTIAL.
FT  DOMAIN 983 1000 CYTOPLASMIC (POTENTIAL).
FT  TRANSMEM 1001 1024 POTENTIAL.
FT  DOMAIN 1025 1047 EXTRACELLULAR (POTENTIAL).
FT  TRANSMEM 1048 1070 POTENTIAL.
FT  DOMAIN 1071 1084 CYTOPLASMIC (POTENTIAL).
FT  TRANSMEM 1085 1102 POTENTIAL.
FT  DOMAIN 1103 1119 EXTRACELLULAR (POTENTIAL).
FT  TRANSMEM 1120 1142 POTENTIAL.
FT  DOMAIN 1143 1187 CYTOPLASMIC (POTENTIAL).
FT  MOD_RES 465 465 PHOSPHORYLATION (BY SIMILARITY).
FT  METAL 880 880 MAGNESIUM (BY SIMILARITY).
FT  METAL 884 884 MAGNESIUM (BY SIMILARITY).
SQ  SEQUENCE 1187 AA; 133112 MW; 3BD185DCDC0C766B0 CRC64;

Query Match 71.4%; Score 30; DB 1; Length 1187;
Best Local Similarity 75.0%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGVTQSRT 8
Db 1112 NGVTSRT 1119

RESULT 8
YDIZ_ECOLI
ID YDIZ_ECOLI STANDARD; PRT; 96 AA.
AC P76207;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOTHETICAL 10.9 KDA PROTEIN IN PFKB-CEDA INTERGENIC REGION.
YDIZ.
GN Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE; 97426617.
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT The complete genome sequence of Escherichia coli K-12.;
RL Science 277:1453-1474(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L15418; AAA17445.1; -.
DR EMBL; U14635; AAC46659.1; ALT_INIT.
DR WORMPEP; C27H5.5; CE06893.
DR INTERPRO; IPR000087; -.
DR PFAM; PF01484; Col_cuticle_N; 1.
DR PFAM; PF01391; Collagen; 3.
KW Cuticle; Connective tissue; Repeat; Multigene family; Collagen.
FT DOMAIN 89 105 TRIPLE-HELICAL REGION.
FT DOMAIN 118 150 TRIPLE-HELICAL REGION.
FT DOMAIN 167 187 TRIPLE-HELICAL REGION.
FT DOMAIN 194 226 TRIPLE-HELICAL REGION.
FT DOMAIN 231 257 TRIPLE-HELICAL REGION.
FT DOMAIN 260 295 TRIPLE-HELICAL REGION.
```

```
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000267; AAC74794.1; -.
DR ECGENE; EG13985; YDIZ.
KW Hypothetical protein.
SQ SEQUENCE 96 AA; 10865 MW; 722C0F845B91A855 CRC64;

Query Match 69.0%; Score 29; DB 1; Length 96;
Best Local Similarity 75.0%; Pred. No. 9.3;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGVTQSRT 8
Db 70 SGLTQSAT 77

RESULT 9
CC36_CAEEL
ID CC36_CAEEL STANDARD; PRT; 307 AA.
AC P34803;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CUTICLE COLLAGEN 36.
GN COL-36 OR C27H5.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE; 94131298.
RA Levy A.D., Kramer J.M.;
RT Identification, sequence and expression patterns of the
RT Caenorhabditis elegans col-36 and col-40 collagen-encoding genes.";
RL Gene 137:281-285(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Pauley A.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NEMATODE CUTICLES ARE COMPOSED LARGELY OF COLLAGEN-LIKE
CC PROTEINS. THE CUTICLE FUNCTIONS BOTH AS AN EXOSKELETON AND AS A
CC BARRIER TO PROTECT THE WORM FROM ITS ENVIRONMENT.
CC -!- SUBUNIT: COLLAGEN POLYPEPTIDE CHAINS ARE COMPLEXED WITHIN THE
CC CUTICLE BY DISULFIDE BONDS AND OTHER TYPES OF COVALENT CROSS-
CC LINKS.
CC -!- SIMILARITY: TO OTHER COLLAGENS. STRONG, TO OTHER CUTICLE
CC COLLAGENS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L15418; AAA17445.1; -.
DR EMBL; U14635; AAC46659.1; ALT_INIT.
DR WORMPEP; C27H5.5; CE06893.
DR INTERPRO; IPR000087; -.
DR PFAM; PF01484; Col_cuticle_N; 1.
DR PFAM; PF01391; Collagen; 3.
KW Cuticle; Connective tissue; Repeat; Multigene family; Collagen.
FT DOMAIN 89 105 TRIPLE-HELICAL REGION.
FT DOMAIN 118 150 TRIPLE-HELICAL REGION.
FT DOMAIN 167 187 TRIPLE-HELICAL REGION.
FT DOMAIN 194 226 TRIPLE-HELICAL REGION.
FT DOMAIN 231 257 TRIPLE-HELICAL REGION.
FT DOMAIN 260 295 TRIPLE-HELICAL REGION.
```



```
FT CONFLICT 774 774 S -> A (IN REF. 2).
SQ SEQUENCE 888 AA; 98251 MW; D54506C5695EC75 CRC64;

Query Match
Best Local Similarity 76.2%; Score 32; DB 1; Length 888;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVTQSRTA 9
    ||| |||
Db 206 GVTTSRTA 213

RESULT 3
POLG_CXB5P STANDARD; PRT; 2185 AA.
AC Q03053;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GENOME POLYPROTEIN (PIC): COAT PROTEIN VP4 (PIA); COAT PROTEIN VP2
DE (PIB); COAT PROTEIN VP3 (PIC); COAT PROTEIN VP1 (PID); PICORNAIN 2A
DE (EC 3.4.22.29) (P2A); CORE PROTEIN P2B; CORE PROTEIN P2C; CORE PROTEIN
DE P3A; GENOME-LINKED PROTEIN VPG (P3B); PICORNAIN 3C (EC 3.4.22.28)
DE (PROTEASE 3C) (P3C); RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (P3D)].
OS Coxsackievirus B5 (strain Peterborough / 1954/UK/85).
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93260398.
RA Zhang G., Willden G., Knowles N.J., McCauley J.W.;
RT "Complete nucleotide sequence of a coxsackie B5 virus and its
RT relationship to swine vesicular disease virus.";
RL J. Gen. Virol. 74:845-853(1993).
CC -!- FUNCTION: P2A AND THE P3C POLYPEPTIDES ARE PROTEASES THAT CLEAVE
CC AT CERTAIN O/G SITES IN THE POLYPROTEIN. THEY ARE CYSTEINE
CC PROTEASES.
CC -!- FUNCTION: IT IS THOUGHT THAT THE P2C PROTEIN ATTACHES TO VESICULAR
CC MEMBRANES AND IS ASSOCIATED WITH VIRAL RNA SYNTHESIS.
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC CLEAVAGE BETWEEN VP4 AND VP2 IS AUTOCATALYTIC; VP1/P2A IS
CC CATALYZED BY P2A; ALL OTHER CLEAVAGES ARE CATALYZED BY P3C.
CC -!- SIMILARITY: P2A PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -!- SIMILARITY: P3C PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; X67706; CAA47944.1; -.
DR PIR; JQ2021; JQ2021.
DR HSSP; P03313; ICOV.
DR INTERPRO; IPR000081; -.
DR INTERPRO; IPR000199; -.
DR INTERPRO; IPR000605; -.
DR INTERPRO; IPR001205; -.
DR INTERPRO; IPR001676; -.
DR INTERPRO; IPR002527; -.
DR PFAM; PF00548; Cys-protease-3C; 1.
DR PFAM; PF00947; Pico_P2A; 1.
DR PFAM; PF01552; Pico_P2B; 1.
DR PFAM; PF00680; RNA_dep_RNA_pol; 1.
DR PFAM; PF00910; RNA_helicase; 1.
DR PFAM; PF00073; rhv; 3.
KW polyprotein; Coat protein; Core protein; Transferase; Myristate;
```

```
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 2 69
FT CHAIN 70 330
FT CHAIN 331 568
FT CHAIN 569 851
FT CHAIN 852 1001
FT CHAIN 1002 1100
FT CHAIN 1101 1429
FT CHAIN 1430 1518
FT CHAIN 1519 1540
FT CHAIN 1541 1723
FT CHAIN 1724 2185
FT LIPID 2 2
FT ACT_SITE 1687 1687
FT ACT_SITE 1701 1701
SQ SEQUENCE 2185 AA; 243298 MW; 3F9EE29F90D59C6F CRC64;

Query Match
Best Local Similarity 76.2%; Score 32; DB 1; Length 2185;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGVTQSRT 8
    :|||:|
Db 836 TGVTSRT 843

RESULT 4
Y08B_BPT4 STANDARD; PRT; 97 AA.
ID Y08B_BPT4
AC P39234;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE HYPOTHETICAL 10.2 KDA PROTEIN IN SEG-GP6 INTERGENIC REGION.
GS Y08B OR 5.4.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae;
OC T4-like phages.
RN [1]
RP SEQUENCE FROM N.A.
RA Marsh R., Mesyanzhinov V.V., Efimov V.P., Prilipov A.G., Duru E.,
RA Song S.;
RL Unpublished observations (xxx-1993).
KW Hypothetical protein.
SQ SEQUENCE 97 AA; 10221 MW; D7C63CD65737C824 CRC64;

Query Match
Best Local Similarity 71.4%; Score 30; DB 1; Length 97;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVTQSRTA 9
    ||| |
Db 58 GVTQPRTS 65

RESULT 5
ARGC_ECOLI STANDARD; PRT; 334 AA.
ID ARGC_ECOLI
AC P11446;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE REDUCTASE (EC 1.2.1.38) (AGPR) (N-
DE ACETYL-GLUTAMATE SEMIALDEHYDE DEHYDROGENASE) (NAGSA DEHYDROGENASE).
GN ARGC.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89121510.
```



```

DR HSSP; P11362; IFGI.
DR MIM; 109135; -.
DR INTERPRO; IPR000719; -.
DR INTERPRO; IPR001245; -.
DR INTERPRO; IPR001777; -.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00041; fn3; 2.
DR PFAM; PF00047; ig; 2.
DR PFAM; PF00069; pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;
KW Transferrase; Phosphorylation; Transmembrane; Signal; Repeat;
KW Immunoglobulin domain; Proto-oncogene; Alternative splicing.
FT SIGNAL 1 18
FT CHAIN 19 887
FT DOMAIN 19 442
FT TRANSMEM 443 465
FT DOMAIN 466 887
FT DOMAIN 29 118
FT DOMAIN 134 205
FT DOMAIN 217 315
FT DOMAIN 319 416
FT DOMAIN 529 800
FT NP_BIND 535 543
FT BINDING 560 560
FT ACT_SITE 665 665
FT MOD_RES 696 696
FT VARSPIC 422 430
FT CONFLICT 296 296
FT CONFLICT 331 331
FT CONFLICT 632 632
FT SEQUENCE 887 AA; 97374 MW; 2478EC1846298EDA CRC64;

Query Match 76.2%; Score 32; DB 1; Length 887;
Best Local Similarity 87.5%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVTSRTA 9
    ||| ||||
Db 205 GVTSRTA 212

RESULT 2
UFO_MOUSE
ID UFO_MOUSE STANDARD; PRT; 888 AA.
AC Q00993;
DT 01-APR-1993 (Rel. 25, Last Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE TYROSINE-PROTEIN KINASE RECEPTOR UFO PRECURSOR (EC 2.7.1.112)
DE (ADHESION-RELATED KINASE).
GN AXL OR UFO OR ARK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-CD-1; TISSUE=HEART;
RX MEDLINE; 92319537.
RA Faust M., Ebensperger C., Schulz A.S., Schleithoff L., Hamelster H.,
RA Bartram C.R., Janssen J.W.G.;
RT "The murine ufo receptor: molecular cloning, chromosomal localization
RL and in situ expression analysis.";
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92019811.
RA Rescigno J., Mansukhani A., Basilico C.;
RT "A putative receptor tyrosine kinase with unique structural
    topology.";

```

Oncogene 6:1909-1913(1991).

-1- FUNCTION: MAY FUNCTION AS A SIGNAL TRANSDUCER BETWEEN SPECIFIC CELL TYPES OF MESODERMAL ORIGIN.

-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

-1- TISSUE SPECIFICITY: IN DISTINCT SUBSTRUCTURES OF A BROAD SPECTRUM OF DEVELOPING TISSUES (IN THE LATE EMBRYOGENESIS). IN CELLS FORMING ORGAN CAPSULES AS WELL AS IN CONNECTIVE TISSUE STRUCTURES (IN ADULT).

-1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC DOMAIN.

-1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

-1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.

-1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 803 ONWARD DUE TO A FRAMESHIFT.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X63535; CAAM5097.1; -

EMBL; X59560; CAAM2158.1; ALT_FRAME.

PIR; S23065; S23065.

PIR; S23251; S23251.

HSSP; P11362; IFGI.

MGD; MGI:1347244; AXL.

INTERPRO; IPR000719; -

INTERPRO; IPR001245; -

INTERPRO; IPR001777; -

INTERPRO; IPR003006; -

PFAM; PF00041; fn3; 2.

PFAM; PF00047; ig; 2.

PFAM; PF00069; pkinase; 1.

PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.

Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding; Transferrase; Phosphorylation; Transmembrane; Signal; Repeat; Immunoglobulin domain.

SIGNAL 1 18

CHAIN 19 888

DOMAIN 19 443

TRANSMEM 444 466

DOMAIN 467 888

DOMAIN 30 119

DOMAIN 135 206

DOMAIN 218 316

DOMAIN 320 417

DOMAIN 530 801

NP_BIND 536 544

BINDING 561 561

ACT_SITE 666 666

MOD_RES 697 697

CARBOHYD 37 37

CARBOHYD 151 151

CARBOHYD 192 192

CARBOHYD 333 333

CARBOHYD 339 339

CARBOHYD 395 395

CARBOHYD 30 45

CONFLICT 69 70

CONFLICT 277 286

CONFLICT 394 394

CONFLICT 486 486

CONFLICT 566 566

CONFLICT 698 699

CONFLICT 728 728

CONFLICT 758 758

CONFLICT 768 768

POTENTIAL.

TYROSINE-PROTEIN KINASE RECEPTOR UFO.

EXTRACELLULAR (POTENTIAL).

POTENTIAL.

CYTOPLASMIC (POTENTIAL).

IG-LIKE DOMAIN.

IG-LIKE DOMAIN.

FIBRONECTIN TYPE-III.

FIBRONECTIN TYPE-III.

PROTEIN KINASE.

ATP (BY SIMILARITY).

BY SIMILARITY.

PHOSPHORYLATION (AUTO-) (BY SIMILARITY).

MISSING (IN SHORT ISOFORM).

P -> T (IN REF. 2).

E -> K (IN REF. 2).

G -> D (IN REF. 2).

97374 MW; 2478EC1846298EDA CRC64;

76.2%; Score 32; DB 1; Length 887;

87.5%; Pred. No. 21;

0; Mismatches 1; Indels 0; Gaps 0;

2 GVTSRTA 9

||| ||||

205 GVTSRTA 212

RESULT 2

UFO_MOUSE

ID UFO_MOUSE STANDARD; PRT; 888 AA.

AC Q00993;

DT 01-APR-1993 (Rel. 25, Last Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE TYROSINE-PROTEIN KINASE RECEPTOR UFO PRECURSOR (EC 2.7.1.112)

DE (ADHESION-RELATED KINASE).

GN AXL OR UFO OR ARK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

RN 11

RP SEQUENCE FROM N.A.

RC STRAIN-CD-1; TISSUE=HEART;

RX MEDLINE; 92319537.

RA Faust M., Ebensperger C., Schulz A.S., Schleithoff L., Hamelster H.,

RA Bartram C.R., Janssen J.W.G.;

RT "The murine ufo receptor: molecular cloning, chromosomal localization

RL and in situ expression analysis.";

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE; 92019811.

RA Rescigno J., Mansukhani A., Basilico C.;

RT "A putative receptor tyrosine kinase with unique structural

RT topology.";

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 09:03:56 ; Search time 58.45 Seconds
(without alignments)
4.919 Million cell updates/sec

Title: US-09-236-468A-2_COPY_496_504
Perfect score: 42
Sequence: 1 SGVTOSRTA 9

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 87993

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	76.2	887	UFO_HUMAN	P30530 homo sapien
2	32	76.2	888	UFO_MOUSE	Q00993 mus musculus
3	32	76.2	2185	POLG_CXB5P	Q03053 c genome po
4	30	71.4	97	Y08B_BP74	P39234 bacterioph
5	30	71.4	334	ARGC_ECOLI	P11446 escherichia
6	30	71.4	515	PEN3_ADEG1	Q64755 avian adeno
7	30	71.4	1187	YH2M_CAEEL	Q27533 caenorhabdi
8	29	69.0	96	YD12_ECOLI	P76207 escherichia
9	29	69.0	307	CC36_CAEEL	P34803 caenorhabdi
10	29	69.0	372	Y412_MYCPN	P75184 mycoplasma
11	29	69.0	393	HVC3_HETFR	P23086 heterodontu
12	29	69.0	430	FTSY_MYCLE	O33010 mycobacteri
13	29	69.0	438	HVCS_HETFR	P23087 heterodontu
14	29	69.0	461	HVCM_HETFR	P23088 heterodontu
15	29	69.0	599	CENB_MOUSE	P27790 mus musculus
16	29	69.0	606	CENB_CRIGR	P48988 cricetulus
17	29	69.0	707	KPC2_CAEEL	P34885 caenorhabdi
18	29	69.0	904	PPRI_YEAST	Q07272 saccharomyc
19	29	69.0	1690	KFLA_HUMAN	Q12756 homo sapien
20	29	69.0	1695	KFLA_MOUSE	P33173 mus musculus
21	29	69.0	1914	STCK_EMENI	Q00706 emericiella
22	29	69.0	2191	POLG_EC06C	Q66474 e genome po
23	29	66.7	70	RPCX_YEAST	P40422 saccharomyc
24	28	66.7	96	VPR_HVIBR	P05928 human immun
25	28	66.7	304	MTDX_HAEIN	P45000 haemophilus
26	28	66.7	451	Y996_SYNY3	P73127 synechocyst
27	28	66.7	512	VG29_BPMU	Q05145 bacterioph
28	28	66.7	539	FIXN_RHIME	Q05572 rhizobium m
29	28	66.7	587	NOT4_YEAST	P34909 saccharomyc
30	28	66.7	611	SNFL_CANGA	Q00372 candida gla
31	28	66.7	622	CZAC_BACTU	Q45743 bacillus th
32	28	66.7	819	YC81_CHLVU	P56370 chlorella v
33	28	66.7	826	HIFA_HUMAN	Q16665 homo sapien

RESULT 1

ID	UFO_HUMAN	STANDARD	PRT	887 AA
AC	P30530;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	TYROSINE-PROTEIN KINASE RECEPTOR UFO PRECURSOR (EC 2.7.1.112) (AXL ONCOGENE).			
DE	ONCOGENE).			
GN	AXL OR UFO.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 92050809.			
RA	Janssen J.W.G., Schulz A.S., Steenvoorden A.C.M., Schmidberger M.,			
RA	Strehl S., Ambros P., Bartram C.R.;			
RT	"A novel putative tyrosine kinase receptor with oncogenic potential.";			
RL	Oncogene 6:2113-2120(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 9201777.			
RA	O'Bryan J.P., Frye R.A., Cogswell P.C., Neubauer A., Kilch B.,			
RA	Prokop C., Espinosa R., le Beau M.M., Earp H., Liu E.T.;			
RT	"AXL, a transforming gene isolated from primary human myeloid			
RT	leukemia cells, encodes a novel receptor tyrosine kinase.";			
RL	Mol. Cell. Biol. 11:5016-5031(1991).			
RN	[3]			
RP	SEQUENCE OF 667-723 FROM N.A.			
RX	MEDLINE; 94067791.			
RA	Lee S.-T., Strunk K.M., Spritz R.A.;			
RT	"A survey of protein tyrosine kinase mRNAs expressed in normal human			
RT	melanocytes.";			
RL	Oncogene 8:3403-3410(1993).			
CC	-!- FUNCTION: MAY FUNCTION AS A SIGNAL TRANSDUCER BETWEEN SPECIFIC			
CC	CELL TYPES OF MESODERMAL ORIGIN.			
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	-!- DISEASE: HAS TRANSFORMING POTENTIAL IN PATIENTS WITH CHRONIC			
CC	MYELOPROLIFERATIVE DISORDER OR CHRONIC MYELOCYTIC LEUKEMIA.			
CC	-!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC			
CC	DOMAIN.			
CC	-!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.			
CC	-!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X57019; CAA40338.1; ALT_INIT.			
DR	EMBL; M76125; AAA61243.1; ALT_INIT.			
DR	EMBL; S65125; AAB20305.1; ALT_INIT.			

Q14432 homo sapien
Q9y2e4 homo sapien
P16604 s genome po
P13900 s genome po
Q03001 homo sapien
Q37409 haemophilus
Q14749 homo sapien
P29731 mastigoclad
P17211 streptomyce
Q29555 sus scrofa
P13255 rattus norv
O59292 pyrococcus

ALIGNMENTS

THIS PAGE BLANK (USPTO)

Search completed: November 8, 2000, 08:53:37
Job time: 368 sec

Query Match 71.4%; Score 30; DB 2; Length 1187;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGVTQSRT 8
:|||||
Db 1112 NGVTSRT 1119

RESULT 14

D64931
hypothetical protein b1724 - Escherichia coli
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 08-Oct-1999
C:Accession: D64931
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: D64931
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-96 <BLAT>
A:Cross-references: GB:AE000267; GB:U000096; NID:gl788011; PIDN:AAC74794.1; PID:gl788018;
A:Experimental source: strain K-12, substrain MG1655

Query Match 69.0%; Score 29; DB 2; Length 96;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGVTQSRT 8
:|||||
Db 70 SGLTQSAT 77

RESULT 15

T37287
collagen 36 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C:Accession: T37287
R:Levy, A.D.; Kramer, J.M.
Gene 137, 281-285, 1993
A:Title: Identification, sequence and expression patterns of the Caenorhabditis elegans
A:Reference number: 221667; MUID:94131298
A:Accession: T37287
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-307 <LEV>
A:Cross-references: EMBL:L15418; PIDN:AAA17445.1
C:Genetics:
A:Gene: col-36
A:Map position: II
A:Introns: 31/2
C:Function:
A:Description: Involved in cuticle assembly
C:Superfamily: unassigned collagens

Query Match 69.0%; Score 29; DB 2; Length 307;
Best Local Similarity 85.7%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGVTQSR 7
:|||||
Db 73 SGVTRSR 79

Query Match 71.4%; Score 30; DB 2; Length 357;
 Best Local Similarity 85.7%; Pred. No. 46;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 GVTQSRRT 8
 Db 294 GVTQDRT 300
 ||||| ||
 ||||| ||

RESULT 9
 F71265
 probable NADH oxidase - syphilis spirochete
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 24-Sep-1999
 C:Accession: F71265
 R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwinn
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
 Science 281, 375-388, 1998
 A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
 A:Reference number: A71250; MUID:98332770
 A:Accession: F71265
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-445 <COL>
 A:Cross-references: GB:AF001260; GB:AE000520; NID:g3323232; PIDN:AA65876.1; PID:g332323
 A:Experimental source: strain Nichols
 C:Genetics:
 A:Gene: TP0921
 C:Superfamily: NADH peroxidase; dihydrolipoamide dehydrogenase homology

Query Match 71.4%; Score 30; DB 2; Length 445;
 Best Local Similarity 66.7%; Pred. No. 57;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 SGVTQSRRTA 9
 Db 339 TGVTPQRAA 347
 :||||| |

RESULT 10
 S52321
 penton base protein - fowl adenovirus 1
 N:Alternate names: envelope protein
 C:Species: Aviadenovirus gall (fowl adenovirus 1, CELO)
 C:Date: 08-May-1995 #sequence_revision 14-Feb-1997 #text_change 26-Aug-1999
 C:Accession: S52321
 R:Akopian, T.A.; Lazareva, S.E.; Tikhomirov, E.E.; Karpov, V.A.; Naroditsky, B.S.
 Submitted to the EMBL Data Library, February 1995
 A:Description: Genes for Avian adenovirus CELO penton base and core polypeptides.
 A:Reference number: S52320
 A:Accession: S52321
 A:Molecule type: DNA
 A:Residues: 1-515 <AKO>
 A:Cross-references: EMBL:Z48167; NID:g755698; PIDN:CAA88101.1; PID:g663165
 C:Superfamily: adenovirus penton protein
 C:Keywords: envelope protein

Query Match 71.4%; Score 30; DB 2; Length 515;
 Best Local Similarity 66.7%; Pred. No. 67;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 SGVTQSRRTA 9
 Db 338 TGVTPQRTA 346
 :||||| ||

RESULT 11
 T08601
 hypothetical protein DG1041 - slime mold (Dictyostelium discoideum)
 C:Species: Dictyostelium discoideum

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 24-Nov-1999
 C:Accession: T08601
 R:Loomis, W.F.; Iranfar, N.
 Submitted to the EMBL Data Library, August 1997
 A:Reference number: Z16451
 A:Accession: T08601
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1021 <LOO>
 A:Cross-references: EMBL:AF019107; NID:g2425053; PID:g2425054
 A:Experimental source: strain AX4
 C:Genetics:
 A:Gene: DG1041
 A:Introns: 15/1
 C:Superfamily: slime mold (Dictyostelium discoideum) hypothetical protein DG1041

Query Match 71.4%; Score 30; DB 2; Length 1021;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 SGVTQSRRTA 9
 Db 496 SGLTQDNTA 504
 ||:|| ||

RESULT 12
 T13107
 tail tip fiber protein gp21 - phage N15
 C:Species: phage N15
 C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 20-Sep-1999
 C:Accession: T13107
 R:Hendrix, R.W.; Ravin, V.K.; Casjens, S.R.; Ford, M.E.; Ravin, N.V.; Smirnov, I.K.
 Submitted to the EMBL Data Library, May 1998
 A:Reference number: Z17603
 A:Accession: T13107
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1061 <HEN>
 A:Cross-references: EMBL:AF064539; NID:g3192683; PID:g3192704; PIDN:AA619057.1
 C:Genetics:
 A:Note: gene 21
 C:Superfamily: phage lambda host specificity protein J

Query Match 71.4%; Score 30; DB 2; Length 1061;
 Best Local Similarity 75.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 SGVTQSRRT 8
 Db 530 SGASQSRRT 537
 ||:||||

RESULT 13
 T26283
 hypothetical protein W08D2.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T26283
 R:Swinburne, J.; Ainscough, R.
 Submitted to the EMBL Data Library, March 1996
 A:Reference number: Z20188
 A:Accession: T26283
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1187 <WIL>
 A:Cross-references: EMBL:Z70271; PIDN:CAA94236.1; GSPDB:GN00022; CESP:W08D2.5
 A:Experimental source: clone W08D2
 C:Genetics:
 A:Gene: CESP:W08D2.5
 A:Map position: 4
 A:Introns: 22/3; 66/2; 83/3; 198/2; 235/2; 261/3; 304/3; 344/3; 379/3; 427/1; 461/3;

R:Zhang, G.; Wilsden, G.; Knowles, N.J.; McCauley, J.W.
J. Gen. Virol. 74, 845-853, 1993
A:Title: Complete nucleotide sequence of a coxsackie B5 virus and its relationship to sw
A:Reference number: JQ2021; MUID:93260398
A:Accession: JQ2021
A:Molecule type: genomic RNA
A:Residues: 1-2185 <ZHA>
A:Cross-references: GB:X67706; NID:g59045; PIDN:CAA7944.1; PID:g59046
C:Superfamily: poliovirus genome polyprotein
C:Keywords: coat protein; core protein; genome-linked protein; nucleotidyltransferase; P
F:1-69/Product: coat protein 1A #status predicted <CIA>
F:70-330/Product: coat protein 1B #status predicted <CIH>
F:331-568/Product: coat protein 1C #status predicted <CIH>
F:569-851/Product: coat protein 1D #status predicted <CID>
F:852-1001/Product: core protein 2A #status predicted <C2A>
F:1002-1100/Product: core protein 2B #status predicted <C2B>
F:1101-1429/Product: core protein 2C #status predicted <C2C>
F:1430-1518/Product: protein 3A #status predicted <P3A>
F:1519-1540/Product: genome-linked protein VPg #status predicted <VPG>
F:1541-1723/Product: proteinase #status predicted <PTS>
F:1724-2185/Product: RNA-directed RNA polymerase #status predicted <RPS>
F:1521/Binding site: phosphoryl-RNA (Tyr) (covalent) #status predicted

Query Match 76.2%; Score 32; DB 1; Length 2185;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGVTQSRT 8
Db 836 TGVTSRT 843

RESULT 6
T23007
hypothetical protein K09C8.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T23007; T23543
R:Kershaw, J.
Submitted to the EMBL Data Library, November 1995
A:Reference number: Z19651
A:Accession: T23007
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1328 <WIL>
A:Cross-references: EMBL:Z68005; PIDN:CAA91994.1; GSPDB:GN000028; CESP:K09C8.5
A:Experimental source: clone F59f3
R:Kershaw, J.
Submitted to the EMBL Data Library, November 1995
A:Reference number: Z19755
A:Accession: T23543
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1328 <W12>
A:Cross-references: EMBL:Z68006; PIDN:CAA91999.1; GSPDB:GN000028; CESP:K09C8.5
A:Experimental source: clone K09C8
C:Genetics:
A:Gene: CESP:K09C8.5
A:Map position: x
A:Introns: 34/1; 85/3; 133/3; 182/2; 220/3; 262/2; 390/3; 442/2; 493/3; 563/2; 586/3; 61

Query Match 73.8%; Score 31; DB 2; Length 1328;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVTQSRTA 9
Db 1151 GVTERTA 1158

RESULT 7

RDECEP
N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) - Escherichia coli
C:Species: Escherichia coli
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 11-Jun-1999
C:Accession: JT0332; A42377; A65203; A30776
R:Parsot, C.; Boyen, A.; Cohen, G.N.; Glansdorff, N.
Gene 68, 273-283, 1988
A:Title: Nucleotide sequence of Escherichia coli argB and argC genes: comparison of N
es.
A:Reference number: JT0331; MUID:89121510
A:Accession: JT0332
A:Molecule type: DNA
A:Residues: 1-334 <PAR>
A:Cross-references: GB:M21446; NID:g145332; PIDN:AAA23477.1; PID:g145333
R:Melnell, T.; Schmitt, E.; Mechulam, Y.; Blanquet, S.
J. Bacteriol. 174, 2323-2331, 1992
A:Title: Structural and biochemical characterization of the Escherichia coli argE gen
A:Reference number: A42377; MUID:92202162
A:Accession: A42377
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-19 <MEI>
A:Cross-references: GB:X55417
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: A65203
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-334 <BLAT>
A:Cross-references: GB:AE000470; GB:U00096; NID:g2367332; PIDN:AAC76940.1; PID:g17903
A:Experimental source: Strain K-12, substrain MG1655
C:Comment: in arginine biosynthesis glutamate is first converted to N-acetylglutamate
e catalyzed by acetylglutamate kinase and N-acetyl-gamma-glutamyl-phosphate reductase
C:Genetics:
A:Gene: argC
A:Map position: 90 min
C:Superfamily: N-acetyl-gamma-glutamyl-phosphate reductase
C:Keywords: arginine biosynthesis; oxidoreductase
F:154/Active site: Cys #status predicted
Query Match 71.4%; Score 30; DB 1; Length 334;
Best Local Similarity 66.7%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 SGVTQSRTA 9
Db 249 SGVTOAQA 257
RESULT 8
S49166
cysteine proteinase (EC 3.4.22.-) precursor - spring vetch
C:Species: Vicia sativa (spring vetch, tare)
C:Date: 16-Feb-1995 #sequence_revision 21-Jul-1995 #text_change 04-Feb-2000
C:Accession: S49166
R:Becker, C.; Senyuk, V.; Shutov, A.; Nong, V.; Fischer, J.; Horstmann, C.; Vaintraub
submitted to the EMBL Data Library, June 1994
A:Description: Molecular characterisation of proteinase A, a papain-like cysteine pro
A:Reference number: S49166
A:Accession: S49166
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-357 <BEC>
A:Cross-references: EMBL:Z34895
C:Superfamily: papain
C:Keywords: cysteine proteinase; hydrolase
F:152,286,307/Active site: Cys, His, Asn #status predicted

C:Keywords: ATP; autophosphorylation; phosphoprotein; transmembrane protein
F:43-113/Domain: immunoglobulin homology <IMM>
F:327-411/Domain: fibronectin type III repeat homology <3FR>
F:528-804/Domain: protein kinase homology <KIN>
F:536-544/Region: protein kinase ATP-binding motif

Query Match 76.2%; Score 32; DB 2; Length 888;
Best Local Similarity 87.5%; Pred. NO. 41;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 GVTQSRTA 9
||| ||||
Db 206 GVTTSRTA 213

RESULT 3
A41527
protein-tyrosine kinase (EC 2.7.1.112) axl precursor, major splice form - human
A:Alternate names: transforming protein axl; UFO receptor
C:Species: Homo sapiens (man)
C:Date: 28-May-1992 #sequence_revision 05-Jan-1996 #text_change 04-Feb-2000
C:Accession: A41527; B38269; I39203; G02782
R:O'Bryan, J.P.; Frye, R.A.; Cogswell, P.C.; Neubauer, A.; Kitch, B.; Prokop, C.; Espinoza, M.; Cell. Biol. 11, 5016-5031, 1991
A:Title: axl, a transforming gene isolated from primary human myeloid leukemia cells, encodes a protein-tyrosine kinase
A:Reference number: A41527; MUID:9201777
A:Accession: A41527
A:Molecule type: mRNA
A:Residues: 1-894 <OAB>
A:Cross-references: GB:M76125
A:Experimental source: axl(+)
A:Accession: B41527
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-337, 'K', 339-428, 438-894 <OAB>
A:Cross-references: GB:M76125; NID:922869; PIDN:AAA61243.1; PID:g292870
A:Experimental source: axl(-)
A:Note: The authors translated the codon AAG for residue 338 as Leu
R:Partanen, J.; Maekela, T.P.; Alitalo, R.; Leivaeslaiho, H.; Alitalo, K.
Proc. Natl. Acad. Sci. U.S.A. 87, 8913-8917, 1990
A:Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.
A:Reference number: A38268; MUID:91062389
A:Accession: B38269
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 674-730 <PAR>
A:Cross-references: GB:M59373; GB:M37781
R:Schulz, A.S.; Schleithoff, L.; Faust, M.; Bartram, C.R.; Janssen, J.W.
Oncogene 8, 509-513, 1993
A:Title: The genomic structure of the human UFO receptor.
A:Reference number: I39203; MUID:93149614
A:Accession: I39203
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-28 <SCH>
A:Cross-references: EMBL:X66030; NID:g37594; PIDN:CAA46829.1; PID:g37595
R:Mitchell, P.J.
submitted to the EMBL Data Library, December 1990
A:Reference number: G09377
A:Accession: G02782
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-302, 'P', 304-429, 'E', 431-638, 'G', 640-894 <MIT>
A:Cross-references: EMBL:X57019; NID:g37592; PIDN:CAA40338.1; PID:g37593
C:Comment: This protein is overexpressed in chronic myelogenous leukemia and induces neo-
C:Genetics:
A:Gene: GDB:AXL; UFO
A:Cross-references: GDB:I33764; OMIM:109135
A:Map position: 19q13.1-19q13.1
C:Function:
A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
C:Superfamily: protein-tyrosine kinase axl; fibronectin type III repeat homology; immunoglobulin homology

C:Keywords: alternative splicing; ATP; autophosphorylation; duplication; glycoprotein
tyrosine-specific protein kinase
F:1-32/Domain: signal sequence #status predicted <SIG>
F:33-894/Product: protein-tyrosine kinase axl, major splice form #status predicted <M
F:49-119/Domain: immunoglobulin homology <IM1>
F:153-207/Domain: immunoglobulin homology <IM2>
F:224-320/Domain: fibronectin type III repeat homology <FN3A>
F:333-417/Domain: fibronectin type III repeat homology <FN3B>
F:448-472/Domain: transmembrane #status predicted <TMM>
F:534-810/Domain: protein kinase homology <KIN>
F:542-550/Region: protein kinase ATP-binding motif
F:43-157,198,339,345,401/Binding site: carboxylate (Asn) (covalent) #status predicted
F:567,585,672/Active site: Lys, Glu, Asp #status predicted
F:677,690/Binding site: magnesium (Asn, Asp) #status predicted
F:703/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi
F:779,821/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status p

Query Match 76.2%; Score 32; DB 1; Length 894;
Best Local Similarity 87.5%; Pred. NO. 41;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 GVTQSRTA 9
||| ||||
Db 212 GVTTSRTA 219

RESULT 4
S23251
protein-tyrosine kinase (EC 2.7.1.112) ark precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jan-2000
C:Accession: S23251
R:Rescigno, J.; Mansukhani, A.; Basilico, C.
Oncogene 6, 1909-1913, 1991
A:Title: A putative receptor tyrosine kinase with unique structural topology.
A:Reference number: S23251; MUID:92019811
A:Accession: S23251
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-942 <RES>
A:Cross-references: EMBL:X59560
C:Genetics:
A:Gene: ark
C:Superfamily: protein-tyrosine kinase axl; fibronectin type III repeat homology; imm
C:Keywords: ATP; autophosphorylation; glycoprotein; phosphoprotein; phosphotransferas
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-942/Product: protein-tyrosine kinase ark #status predicted <MAT>
F:147-201/Domain: immunoglobulin homology <IMM>
F:327-411/Domain: fibronectin type III repeat homology <3FR>
F:446-466/Domain: transmembrane #status predicted <TMM>
F:528-804/Domain: protein kinase homology <KIN>
F:536-544/Region: protein kinase ATP-binding motif
F:151,192,333,339,395,677/Binding site: carboxylate (Asn) (covalent) #status predict

Query Match 76.2%; Score 32; DB 2; Length 942;
Best Local Similarity 87.5%; Pred. NO. 44;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 GVTQSRTA 9
||| ||||
Db 206 GVTTSRTA 213

RESULT 5
J02021
genome polyprotein - coxsackievirus B5 (strain 1954/UK/85)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core
polymerase (EC 2.7.7.48)
C:Species: coxsackievirus B5
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 16-Jul-1999
C:Accession: J02021

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:53:36 ; Search time 99.87 Seconds
(without alignments)
5.719 Million cell updates/sec

Title: US-09-236-468A-2_COPY_496_504

Perfect score: 42

Sequence: 1 SGVTQSRTA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 182106

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_65:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	33	78.6	261	2 J01024	hypothetical 30K p
2	32	76.2	888	2 S23065	ufo protein - mous
3	32	76.2	894	1 A41527	protein-tyrosine k
4	32	76.2	942	2 S23251	protein-tyrosine k
5	32	76.2	2185	1 J02021	genome polyprotein
6	31	73.8	1328	2 T23007	hypothetical prote
7	30	71.4	334	1 R0CEP	N-acetyl-gamma-glu
8	30	71.4	357	2 S49166	cysteine proteinas
9	30	71.4	445	2 F71265	probable NADH oxid
10	30	71.4	515	2 S23221	penton base protei
11	30	71.4	1021	2 T08601	hypothetical prote
12	30	71.4	1061	2 T13107	tail tip fiber pro
13	30	71.4	1187	2 T26283	hypothetical prote
14	29	69.0	96	2 D64931	hypothetical prote
15	29	69.0	307	2 T37287	collagen 36 - Caen
16	29	69.0	385	2 S73557	MG412 homolog C12
17	29	69.0	393	1 HVRKCL	Ig mu chain C regi
18	29	69.0	421	2 H69490	formylmethanofuran
19	29	69.0	438	1 HVRKCS	Ig mu chain C regi
20	29	69.0	461	1 HVRKCO	Ig mu chain C regi
21	29	69.0	547	2 T22856	hypothetical prote
22	29	69.0	606	2 S70358	centromere protein
23	29	69.0	707	1 A53530	protein kinase C (
24	29	69.0	904	1 RGBYPI	regulatory protein
25	29	69.0	1110	2 T19673	hypothetical prote
26	29	69.0	1332	2 T15670	hypothetical prote
27	29	69.0	1440	2 T27942	lin-15B protein -
28	29	69.0	1695	2 A56921	kinesin family pro
29	28	66.7	32	2 S21547	T-cell receptor al

30	28	66.7	70	2 S58932	DNA-directed RNA p
31	28	66.7	113	2 S26266	T-cell receptor be
32	28	66.7	113	2 PT0741	T-cell receptor be
33	28	66.7	113	2 S17385	T-cell receptor be
34	28	66.7	113	2 S17386	T-cell receptor be
35	28	66.7	113	2 S26262	T-cell receptor be
36	28	66.7	113	2 S26263	T-cell receptor va
37	28	66.7	115	2 JC5908	T-cell receptor va
38	28	66.7	146	2 S26408	T-cell receptor be
39	28	66.7	152	2 S21826	T-cell receptor be
40	28	66.7	304	2 C64109	site-specific DNA-
41	28	66.7	336	2 T23902	hypothetical prote
42	28	66.7	424	2 T49317	probable 26s prote
43	28	66.7	451	1 S75239	hypothetical prote
44	28	66.7	480	2 T20154	hypothetical prote
45	28	66.7	539	1 S39988	cytochrome-c oxida

ALIGNMENTS

RESULT 1

J01024

hypothetical 30K protein (DmRP140 5' region) - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 20-Sep-1999

C:Accession: J01024; JS0573

R:Sitler, S.; Oldenburg, I.; Petersen, G.; Bautz, E.K.F.

Gene 100, 155-162, 1991

A:Title: Analysis of the promoter region of the housekeeping gene DmRP140 by sequence

A:Reference number: PQ0154; MUID:91276237

A:Accession: J01024

A:Molecule type: DNA

A:Residues: 1-261 <SIT>

A:Cross-references: GB:M62975

A:Accession: JS0573

A:Molecule type: mRNA

A:Residues: 1-261 <SI2>

A>Note: the authors translated the codon TTC for residue 64 as Ser

C:Genetics:

A:Gene: FlyBase:Rp1140

A:Cross-references: FlyBase:FBgn0003276

A:Introns: 58/2; 113/3; 143/2

C:Superfamily: Drosophila hypothetical 30K protein (DmRP140 5' region)

Query Match 78.6%; Score 33; DB 2; Length 261;
Best Local Similarity 87.5%; Pred. No. 7;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVTQSRTA 9

DB 86 GVTQSRVA 93

RESULT 2

S23065

ufo protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 19-Feb-1994 #sequence_revision 01-Sep-1995 #text_change 21-Jan-2000

C:Accession: S23065

R:Faust, M.; Ebensperger, C.; Schulz, A.S.; Schleithoff, L.; Hameister, H.; Bartram,

Oncogene 7, 1287-1293, 1992

A:Title: The murine ufo receptor: molecular cloning, chromosomal localization and in

A:Reference number: S23065; MUID:92319537

A:Accession: S23065

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-888 <FAU>

A:Cross-references: EMBL:X63535; NID:g55121; PIDN:CAA45097.1; PID:g55122

C:Genetics:

A:Gene: ufo

C:Superfamily: protein-tyrosine kinase axl; fibronectin type III repeat homology; imm

THIS PAGE BLANK (USPTO)

; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: AFOL-0136
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US94-02191-10

Query Match 66.7%; Score 28; DB 4; Length 16;
Best Local Similarity 75.0%; Pred. No. 3;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVTQSRTA 9
| | | | |
Db 2 GVTQORRA 9

RESULT 15
US-08-301-915-1
; Sequence 1, Application US/08301915
; Patent No. 5861161
; GENERAL INFORMATION:
; APPLICANT: COHEN, Eric A.
; APPLICANT: BERGERON, Dominique
; APPLICANT: CHECROUNE, Florent
; APPLICANT: YAO, Xiao-Jian
; APPLICANT: PIGNAC-KOBINGER, Gary
; TITLE OF INVENTION: PROTEIN TARGETING INTO HIV VIRIONS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KLAUBER & JACKSON
; STREET: Continental Plaza, 411 Hackensack Avenue
; CITY: Hackensack
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/301,915
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: JACKSON, David A.
; REGISTRATION NUMBER: 26,742
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-467-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-301-915-1

Query Match 66.7%; Score 28; DB 2; Length 96;

Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2 GVTQSRTA 9
| | | | |
Db 82 GVTQORRA 89
Search completed: November 8, 2000, 08:49:24
Job time: 117 sec

```

; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genetech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,461
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/170558
; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasek, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 854C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 894 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-445-461-34

```

```

Query Match 76.2%; Score 32; DB 3; Length 894;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

```

```

Qy 2 GVTQSRTA 9
   ||| ||||
Db 212 GVTTSRTA 219

```

```

RESULT 13
US-08-707-237A-84
; Sequence 84, Application US/08/07237A
; Patent No. 5830713
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Capello, Joseph
; APPLICANT: Crissman, John W.
; APPLICANT: Dorman, Mary A.
; TITLE OF INVENTION: METHODS FOR PREPARING SYNTHETIC
; TITLE OF INVENTION: REPETITIVE DNA
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/707,237A
; FILING DATE: 03-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,155
; FILING DATE: 29-DEC-1993

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/053,049
; FILING DATE: 22-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/609,716
; FILING DATE: 06-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/269,429
; FILING DATE: 09-NOV-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/114,618
; FILING DATE: 29-OCT-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/927,258
; FILING DATE: 04-NOV-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-55186-10/WHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 761 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-707-237A-84

```

```

Query Match 69.0%; Score 29; DB 2; Length 761;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 2 GVTQSRTA 9
   ||||| |
Db 15 GVTQNRLA 22

```

```

RESULT 14
PCT-US94-02191-10
; Sequence 10, Application PC/TUS9402191
; GENERAL INFORMATION:
; APPLICANT: Weiner, David B
; APPLICANT: Levy, David
; APPLICANT: Refaelli, Yosef
; TITLE OF INVENTION: VPR Function and Activity
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: United States of America
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 mbMD
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/02191
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/019,601
; FILING DATE: 19-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/167,608
; FILING DATE: 15-DEC-1993

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 894 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-445-640-34

Query Match 76.2%; Score 32; DB 1; Length 894;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVTQSRTA 9
||| ||||
DB 212 GVTTSRTA 219

RESULT 10
US-08-170-558-34
; Sequence 34, Application US/08170558
; Patent No. 6001621
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie R.
; APPLICANT: Scadden, David T.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Protein Tyrosine Kinases
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/170,558
; FILING DATE: 20-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 854C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 894 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-170-558-34

Query Match 76.2%; Score 32; DB 3; Length 894;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVTQSRTA 9
||| ||||
DB 212 GVTTSRTA 219

RESULT 11

US-08-447-314-34
; Sequence 34, Application US/08447314
; Patent No. 6087144
; GENERAL INFORMATION:
; APPLICANT: Scadden, David T.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Protein Tyrosine Kinases
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,314
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/170558
; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 854C1D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 894 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-447-314-34

Query Match 76.2%; Score 32; DB 3; Length 894;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVTQSRTA 9
||| ||||
DB 212 GVTTSRTA 219

RESULT 12
US-08-445-461-34
; Sequence 34, Application US/08445461
; Patent No. 6096527
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie R.
; APPLICANT: Scadden, David T.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Protein Tyrosine Kinases
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California

;; TITLE OF INVENTION: Protein Tyrosine Kinases
;; NUMBER OF SEQUENCES: 35
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Genentech, Inc.
;; STREET: 460 Point San Bruno Blvd
;; CITY: South San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94080
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: patin (Genentech)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/445,461
;; FILING DATE: 22-MAY-1995
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/170558
;; FILING DATE: 20-DEC-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/157563
;; FILING DATE: 23-NOV-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hasak, Janet E.
;; REGISTRATION NUMBER: 28,616
;; REFERENCE/DOCKET NUMBER: 854C3
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415/225-1896
;; TELEFAX: 415/952-9881
;; TELEX: 910/371-7168
;; INFORMATION FOR SEQ ID NO: 35:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 888 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; US-08-445-461-35

Query Match 76.2% Score 32; DB 3; Length 888;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVTQSRTA 9
||| ||||
Db 206 GVTTSRTA 213

RESULT 8
US-08-372-892-2
;; Sequence 2, Application US/08372892
;; Patent No. 5468634
;; GENERAL INFORMATION:
;; APPLICANT: Liu, Edison T.
;; TITLE OF INVENTION: AXL Oncogene
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Kenneth D. Sibley; Bell, Seltzer, Park and
;; ADDRESSEE: Gibson
;; STREET: Post Office Drawer 34009
;; CITY: Charlotte
;; STATE: No. 5468634th Carolina
;; COUNTRY: U.S.A.
;; ZIP: 28234
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/372,892
;; FILING DATE:

;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/718,572
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Sibley, Kenneth D.
;; REGISTRATION NUMBER: 31,665
;; REFERENCE/DOCKET NUMBER: 5470-15
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 919-881-3140
;; TELEFAX: 919-881-3175
;; TELEX: 575102
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 894 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-372-892-2

Query Match 76.2% Score 32; DB 1; Length 894;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVTQSRTA 9
||| ||||
Db 212 GVTTSRTA 219

RESULT 9
US-08-445-640-34
;; Sequence 34, Application US/08445640
;; Patent No. 5709858
;; GENERAL INFORMATION:
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Mark, Melanie R.
;; APPLICANT: Scadden, David T.
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Baron, Will F.
;; TITLE OF INVENTION: Protein Tyrosine Kinases
;; NUMBER OF SEQUENCES: 35
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Genentech, Inc.
;; STREET: 460 Point San Bruno Blvd
;; CITY: South San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94080
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: patin (Genentech)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/445,640
;; FILING DATE: 22-MAY-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/170558
;; FILING DATE: 20-DEC-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/157563
;; FILING DATE: 23-NOV-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hasak, Janet E.
;; REGISTRATION NUMBER: 28,616
;; REFERENCE/DOCKET NUMBER: 854C2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415/225-1896
;; TELEFAX: 415/952-9881
;; TELEX: 910/371-7168
;; INFORMATION FOR SEQ ID NO: 34:

; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 854C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 888 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-445-640-35

Query Match 76.2%; Score 32; DB 1; Length 888;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVTQSRTA 9
Db 206 GVTTSRTA 213

RESULT 5
US-08-170-558-35
; Sequence 35, Application US/08170558
; Patent No. 6001621
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie R.
; APPLICANT: Scadden, David T.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Protein Tyrosine Kinases
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/170,558
; FILING DATE: 20-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 854C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 888 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-170-558-35

Query Match 76.2%; Score 32; DB 3; Length 888;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVTQSRTA 9
Db 206 GVTTSRTA 213

RESULT 6
US-08-447-314-35
; Sequence 35, Application US/08447314
; Patent No. 6087144
; GENERAL INFORMATION:
; APPLICANT: Scadden, David T.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Protein Tyrosine Kinases
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,314
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/170558
; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 854C1D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 888 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-447-314-35

Query Match 76.2%; Score 32; DB 3; Length 888;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVTQSRTA 9
Db 206 GVTTSRTA 213

RESULT 7
US-08-445-461-35
; Sequence 35, Application US/08445461
; Patent No. 6096527
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie R.
; APPLICANT: Scadden, David T.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.

Db 496 SGVTSRSTA 504

RESULT 2
PCT-US95-07085-2
; Sequence 2, Application PC/TUS9507085
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone Receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07085
; FILING DATE: 05-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-393
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-07085-2

Query Match 100.0%; Score 42; DB 4; Length 541;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SGVTSRSTA 9
|||||
Db 496 SGVTSRSTA 504

RESULT 3
US-08-372-892-4
; Sequence 4, Application US/08372892
; Patent No. 5468634
; GENERAL INFORMATION:
; APPLICANT: Liu, Edison T.
; TITLE OF INVENTION: AXL Oncogene
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley; Bell, Seltzer, Park and
; ADDRESSEE: Gibson
; STREET: Post Office Drawer 34009
; CITY: Charlotte
; STATE: No. 5468634th Carolina
; COUNTRY: U.S.A.
; ZIP: 28234
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/372,892
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/718,572
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-15.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 885 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-372-892-4

Query Match 76.2%; Score 32; DB 1; Length 885;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVTQSRSTA 9
|||||
Db 212 GVTTSRTA 219

RESULT 4
US-08-445-640-35
; Sequence 35, Application US/08445640
; Patent No. 5709858
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie R.
; APPLICANT: Scadden, David T.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Protein Tyrosine Kinases
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,640
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/170558
; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:49:23 ; Search time 97.15 Seconds
(without alignments)
1.553 Million cell updates/sec

Title: us-09-236-468a-2_copy_496_504

Perfect score: 42

Sequence: 1 SGVTQSRKA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 164575

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	42	100.0	541	3	US-08-468-011A-2
2	42	100.0	541	4	PCT-US95-07085-2
3	32	76.2	888	1	US-08-372-892-4
4	32	76.2	888	1	US-08-445-640-35
5	32	76.2	888	3	US-08-170-558-35
6	32	76.2	888	3	US-08-447-314-35
7	32	76.2	888	3	US-08-445-461-35
8	32	76.2	894	1	US-08-372-892-2
9	32	76.2	894	1	US-08-445-640-34
10	32	76.2	894	3	US-08-170-558-34
11	32	76.2	894	3	US-08-447-314-34
12	32	76.2	894	3	US-08-445-461-34
13	29	69.0	761	2	US-08-707-237A-84
14	28	66.7	16	4	PCT-US94-02191-10
15	28	66.7	96	2	US-08-301-915-1
16	28	66.7	96	3	US-08-524-694A-1
17	28	66.7	322	1	US-08-327-494A-2
18	28	66.7	322	4	PCT-US95-13659-2
19	28	66.7	623	3	US-09-041-991A-6
20	28	66.7	805	2	US-08-480-473B-4
21	28	66.7	805	3	US-08-915-213-4
22	28	66.7	805	4	PCT-US96-10251-4
23	28	66.7	826	1	US-08-785-241-6
24	28	66.7	826	2	US-08-480-473B-2
25	28	66.7	826	3	US-08-915-213-2
26	28	66.7	826	4	PCT-US96-10251-2
27	28	66.7	943	2	US-08-808-982-7
28	27	64.3	16	2	US-08-637-759B-52

Sequence 52, Appl
Sequence 88, Appl
Sequence 27, Appl
Sequence 27, Appl
Sequence 6, Appl
Sequence 4, Appl
Sequence 3, Appl
Sequence 1, Appl
Sequence 2, Appl
Sequence 4, Appl
Sequence 59, Appl
Sequence 2, Appl
Sequence 4, Appl
Sequence 73, Appl
Sequence 18, Appl
Sequence 1, Appl
Sequence 2, Appl

29 27 64.3 16 3 US-08-871-355A-52
30 27 64.3 32 3 US-08-484-322-88
31 27 64.3 92 1 US-08-181-492B-27
32 27 64.3 92 4 PCT-US95-00408-27
33 27 64.3 100 1 US-08-681-812-6
34 27 64.3 193 2 US-08-861-549-4
35 27 64.3 205 2 US-08-861-549-3
36 27 64.3 205 2 US-08-861-549-3
37 27 64.3 293 2 US-08-489-141A-2
38 27 64.3 293 2 US-08-907-492A-4
39 27 64.3 293 2 US-08-576-626A-59
40 27 64.3 295 2 US-08-907-492A-2
41 27 64.3 376 2 US-08-709-923-4
42 27 64.3 420 2 US-08-846-762-73
43 27 64.3 533 3 US-08-726-214-18
44 27 64.3 639 2 US-08-637-899-1
45 27 64.3 1134 3 US-08-726-214-2

ALIGNMENTS

RESULT 1
US-08-468-011A-2
; Sequence 2, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESS: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-011A-2

Query Match 100.0%; Score 42; DB 3; Length 541;
Best Local Similarity 100.0%; Pred. No. 0.11; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;
Qy 1 SGVTQSRKA 9
|||||||

DR N-PSDB; Q28179.
XX
PT Variable regions of b-chain of T-lymphocyte receptors and their
PT DNA - useful as immuno:modulant(s) and for diagnosing immune
PT disorders
XX
XX
PS Claim 7; Page 43; 75pp; French.
XX
CC RNA was isolated from peripheral lymphocytes and converted to cDNA
CC using a C-beta-specific primer. The cDNA was amplified by anchored
CC PCR using C-beta and polyC primers, then amplified again using a
CC different C-beta specific primer. The amplified product was SacII-
CC restricted, inserted into Bluescript SK+ vector and used to transform
CC E.coli XL-Iblue. Transformants were screened with a C-beta specific
CC probe and DNA from positive clones was sequenced in the C-beta
CC region. The sequence designated "IGR b 08" corresponds to a 5'
CC extension of the clones Vb12A1 and PL25 which are two of the
CC previously identified members of the V beta 5 subfamily. The
CC peptide encoded by it can be used to block T cell epitopes and in
CC vaccines. See also Q28173-Q28228.
XX
SQ Sequence 113 AA;

Query Match 66.7%; Score 28; DB 13; Length 113;
Best Local Similarity 75.0%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SGVTQSRT 8
Db 21 agvtqspt 28
:|||||I

Search completed: November 8, 2000, 08:51:52
Job time: 264 sec

```

XX HIV-1; HIV-2; virion; Vpr; Vpx; p6; chimeric; infection; anti-HIV;
KW gene therapy.
XX
OS Human immunodeficiency virus type 1.
XX
PN US6043081-A.
XX
PD 28-MAR-2000.
XX
PF 07-SEP-1995; 95US-0524694.
XX
PR 07-SEP-1994; 94US-0301915.
XX
PA (UYMO-) UNIV MONTREAL.
XX
PI Yao X, Pignac-Kobinger G, Checrone F, Cohen EA, Bergeron D;
XX WPI; 2000-270343/23.
DR
XX Expression vector useful for reducing infectivity of HIV or for
XX targeting into HIV virions, comprises nucleic acid segment encoding
PT recombinant or chimeric protein comprising a Vpr/Vpx virion
PT incorporation domain.
XX
XX Claim 1; Column 7-8; 32pp; English.
PS
XX The present invention describes an expression vector (I) comprising a
XX nucleic acid segment encoding a recombinant protein for interfering with
XX the incorporation of native Vpr/Vpx into HIV-1 or HIV-2 virion or a
XX chimeric protein that is incorporated into an HIV-1 or HIV-2 virion,
XX operably linked to a promoter. Also described are: (1) an isolated
XX eukaryotic or prokaryotic cell transformed with (I); (2) a composition
XX for reducing infectivity of HIV-1 or HIV-2 in vitro comprising an
XX effective amount of (I) in association with a pharmaceutically
XX acceptable carrier; and (3) a composition for targeting into an HIV-1 or
XX HIV-2 virion comprising an effective amount of (I) in association with a
XX pharmaceutically acceptable carrier. (I) has anti-HIV activity and can
XX be used in gene therapy. (I) is useful for reducing infectivity of HIV
XX in vitro and for targeting into an HIV-1 or HIV-2 virion. The present
XX sequence represents the Vpr protein from an HIV-1 strain, which is used
XX in the exemplification of the present invention.
XX
SQ Sequence 96 AA;

Query Match 66.7%; Score 28; DB 21; Length 96;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0.

QY 2 GVTQSRTA 9
DB 82 gvtqrra 89

RESULT 14
R26967
ID R26967 standard; Protein; 113 AA.
XX
AC R26967;
XX
XX 11-FEB-1993 (first entry)
XX
DE Human T lymphocyte receptor V-beta 5 subfamily segment.
XX
XX TCR; IGR b 09; variable region; immunomodulation;
KW polymerase chain reaction; T cell receptor.
XX
XX Homo sapiens.
OS
XX WO9213950-A.
PN
XX 20-AUG-1992.
PD

```

```

XX 12-FEB-1992; 92WO-FR00130.
XX
XX 12-FEB-1991; 91FR-0001613.
PR 12-APR-1991; 91FR-0004523.
XX
XX (ROUS ) ROUSSEL-UCLAF.
XX
XX Ferradini L, Hercend T, Roman-Roman S, Triebel F;
XX WPI; 1992-300036/36.
DR N-PSDB; Q28180.
XX
XX Variable regions of b-chain of T-lymphocyte receptors and their
PT DNA - useful as immuno:modulant(s) and for diagnosing immune
PT disorders
XX
XX Claim 7; Page 44; 75pp; French.
XX
XX RNA was isolated from peripheral lymphocytes and converted to cDNA
XX using a C-beta-specific primer. The cDNA was amplified by anchored
XX PCR using C-beta and polyC primers, then amplified again using a
XX different C-beta specific primer. The amplified product was SacI-
XX restricted, inserted into Bluescript SK+ vector and used to transform
XX E.coli XL-blue. Transformants were screened with a C-beta specific
XX probe and DNA from positive clones was sequenced in the C-beta
XX region. The sequence designated "IGR b 09" corresponds to a 5'
XX extension of the clones VB12A1 and PL25 which are two of the
XX previously identified members of the V beta 5 subfamily. The
XX peptide encoded by it can be used to block T cell epitopes and in
XX vaccines. See also Q28173-Q28228.
XX
SQ Sequence 113 AA;

Query Match 66.7%; Score 28; DB 13; Length 113;
Best Local Similarity 75.0%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0.

QY 1 SCVTSQRT 8
DB 21 agvtqsgpt 28

RESULT 15
R26966
ID R26966 standard; Protein; 113 AA.
XX
AC R26966;
XX
XX 11-FEB-1993 (first entry)
XX
DE Human T lymphocyte receptor V-beta 5 subfamily segment.
XX
XX TCR; IGR b 08; variable region; immunomodulation;
KW polymerase chain reaction; T cell receptor.
XX
XX Homo sapiens.
OS
XX WO9213950-A.
PN
XX 20-AUG-1992.
PD
XX
XX 12-FEB-1992; 92WO-FR00130.
XX
XX 12-FEB-1991; 91FR-0001613.
PR 12-APR-1991; 91FR-0004523.
XX
XX (ROUS ) ROUSSEL-UCLAF.
XX
XX Ferradini L, Hercend T, Roman-Roman S, Triebel F;
XX WPI; 1992-300036/36.
DR

```

CC inhibitors or to introduce and remove double bonds in synthetic and
 CC naturally occurring polymeric ring systems (Delta 7-sterol
 CC reductase catalyses the conversion of 7-dehydrocholesterol to
 CC cholesterol)).
 XX
 SQ Sequence 475 AA;

Query Match 69.0%; Score 29; DB 20; Length 475;
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVTQSRTA 9
 |||||
 Db 17 gvtndrta 24

RESULT 11

R10175
 ID R10175 standard; Protein; 96 AA.

XX AC R10175;

XX 21-MAR-1991 (first entry)

DE Rap (R) protein encoded by HIV strain BRU vpr gene.

XX Viral protein R gene; vpr; rap gene; trans-activation; CAT; HIV;
 KW ELI; MAL; BRV.

XX Human immunodeficiency virus.

XX WO9015875-A.

XX 27-DEC-1990.

XX 01-JUN-1990; 90WO-US03126.

XX 02-JUN-1989; 89US-0361028.

XX (DANA-) DANA FARBER CANCER.

PI Haseltine WA, Terwilliger E, Cohen E;

XX WPI; 1991-022239/03.

XX Vector contg. active viral protein R gene encoding protein -
 PT having ability to trans-activate, useful for producing high
 PT levels of desired heterologous gene prod.

XX Disclosure; fig 5; 39pp; English.

XX This viral protein R (rap) is encoded by a vpr gene, from the HIV
 CC strain BRU, contained in a vector. The R protein is capable of
 CC trans-activating a desired heterologous prod. e.g. chloramphenicol
 CC acetyl transferase (CAT) in a wide range of expression systems.
 CC The HIV strains MAL and BRV can also be used as sources of the HIV
 CC active gene (vpr). However strains HXBc2, BH5 and BH10 do not
 CC encode fully functional vpr proteins. See also R1074.

XX Sequence 96 AA;

Query Match 66.7%; Score 28; DB 12; Length 96;
 Best Local Similarity 75.0%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVTQSRTA 9
 |||||
 Db 82 gvtqgrta 89

RESULT 12

DE HIV-1 LAI strain Vpr protein SEQ ID NO:1.

R94544

ID R94544 standard; protein; 96 AA.

XX AC R94544;

XX 17-OCT-1996 (first entry)

DE Native Vpr protein from HIV-1 LAI strain isolate.

XX Native; Vpr protein; HIV-1 LAI strain isolate; chimaeric molecule;
 KW chimeric; infection; replication; reduction; RNase; protease;
 KW virion assembly; morphogenesis; steric hindrance inducer;
 KW targeted; mature virion; structural organisation; gene therapy;
 KW functional integrity.

XX Human immunodeficiency virus type 1.

XX WO9607741-A1.

PD 14-MAR-1996.

XX 07-SEP-1995; 95WO-CA00510.

PR 07-SEP-1994; 94US-0301915.

XX (UYMO-) UNIV MONTREAL.

XX Bergeron D, Checroune F, Cohen EA, Pignac-Kobinger G;

PI Yao X.

XX WPI; 1996-171615/17.

XX Targeting mature HIV virion(s) using HIV Vpr, Vpx or p6 proteins -
 PT for affecting structural organisation or function or for inhibiting
 PT replication

XX Claim 8; Page 43; 68pp; English.

XX The present sequence is the native Vpr protein from the HIV-1 LAI
 CC strain isolate, which can be used in the development of a chimaeric
 CC mol. comprising the present sequence and another mol., pref. a HIV
 CC infectivity or replication reducing protein fragment, i.e. a RNase
 CC and/or protease, a virion assembly and/or morphogenesis steric
 CC hindrance inducer and/or an effector of a viral protein
 CC interaction responsible for viral infectivity and/or replication.
 CC The chimaeric mol. can be specifically targeted into the mature
 CC HIV-1 virion, to affect its structural organisation and/or
 CC functional integrity, i.e. gene therapy of HIV-1 infection. Vpr
 CC protein fragments can also be used to prevent viral replication by
 CC interfering with the protein interactions responsible for Vpr
 CC incorporation into the mature HIV-1 genome.

XX Sequence 96 AA;

Query Match 66.7%; Score 28; DB 17; Length 96;
 Best Local Similarity 75.0%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVTQSRTA 9
 |||||
 Db 82 gvtqgrta 89

RESULT 13

Y53247
 ID Y53247 standard; protein; 96 AA.

XX AC Y53247;

XX 17-JUL-2000 (first entry)

DE HIV-1 LAI strain Vpr protein SEQ ID NO:1.

PR 20-JAN-1999; 99US-0235609.
PR 22-JAN-1998; 98US-0072134.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Agostino MJ, Clark HF, Collins-Racie LA, Fechtel K;
PI Jacobs K, Lavallie ER, McCoy JM, Merberg D, Steininger RJ;
PI Treacy M, Wong GG;
XX
DR WPI; 1999-458682/38.
DR N-PSDB; X90448.
XX
XX New polynucleotides encoding secreted human proteins derived from,
PT e.g. fetal brain potentially used as immunostimulators
XX
PS Claim 34; Page 126-128; 139pp; English.
XX
CC The present sequence represents a human secreted protein. Human secreted
CC protein polynucleotides and proteins are predicted to have biological
CC activities which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals, although no
CC supporting data is given. Suggested activities include nutritional
CC activity, cytokine and cell proliferation/differentiation activity,
CC immune stimulating (e.g. as vaccines) or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, and tumour
CC inhibition activity. The polynucleotides are also stated to be useful
CC for gene therapy.
XX
SQ Sequence 475 AA;

Query Match 69.0%; Score 29; DB 20; Length 475;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVTOSRTA 9
Db 17 gvtndrta 24

RESULT 9
W93572
ID W93572 standard; Protein; 475 AA.
XX
AC W93572;
XX
DT 17-JUN-1999 (first entry)
XX
DE Human Delta7-sterol reductase protein.
XX
KW Human; Delta7-sterol reductase; hereditary; Smith-Lemli-Opitz syndrome;
KW diagnosis; screening; double bond removal; 7-dehydrocholesterol;
KW organic polymeric ring; cholesterol.
XX
OS Homo sapiens.
XX
PN DE19739940-A1.
XX
PD 18-MAR-1999.
XX
PF 11-SEP-1997; 97DE-1039940.
XX
DE Human Delta7-sterol reductase protein.
XX
KW Human; Delta7-sterol reductase; hereditary; Smith-Lemli-Opitz syndrome;
KW diagnosis; screening; double bond removal; 7-dehydrocholesterol;
KW organic polymeric ring; cholesterol.
XX
XX Homo sapiens.
XX
PN DE19739940-A1.
XX
PD 18-MAR-1999.
XX
PF 11-SEP-1997; 97DE-1039940.
XX
PR 11-SEP-1997; 97DE-1039940.
XX
PA (GLOS/) GLOSSMANN H.
XX
PI Fitzky B, Glossmann H, Moebius F;
XX
DR WPI; 1999-191430/17.
DR N-PSDB; X23386.

XX Human Delta7-sterol reductase polypeptide - useful for diagnosis or
PT treatment of genetic defects e.g. hereditary Smith-Lemli-Opitz
PT syndrome
XX
PS Claim 1; Page 27-28; 62pp; German.
XX
CC This invention describes a human Delta7-sterol reductase. The encoding
CC DNA can be used to diagnose or correct human Delta7-sterol reductase gene
CC defects e.g. hereditary Smith-Lemli-Opitz syndrome, and to produce the
CC recombinant Delta7-sterol reductase polypeptide, which can be used to
CC replace a defective Delta7-sterol reductase enzyme in humans or other
CC animals. It is also useful to screen for Delta7-sterol reductase
CC inhibitors or to introduce and remove double bonds in synthetic and
CC naturally occurring organic polymeric ring systems (Delta 7-sterol
CC reductase catalyses the conversion of 7-dehydrocholesterol to
CC cholesterol).
XX
SQ Sequence 475 AA;

Query Match 69.0%; Score 29; DB 20; Length 475;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVTOSRTA 9
Db 17 gvtndrta 24

RESULT 10
W93573
ID W93573 standard; Protein; 475 AA.
XX
AC W93573;
XX
DT 17-JUN-1999 (first entry)
XX
DE Human Delta7-sterol reductase protein.
XX
KW Human; Delta7-sterol reductase; hereditary; Smith-Lemli-Opitz syndrome;
KW diagnosis; screening; double bond removal; 7-dehydrocholesterol;
KW organic polymeric ring; cholesterol.
XX
OS Homo sapiens.
XX
PN DE19739940-A1.
XX
PD 18-MAR-1999.
XX
PF 11-SEP-1997; 97DE-1039940.
XX
PR 11-SEP-1997; 97DE-1039940.
XX
PA (GLOS/) GLOSSMANN H.
XX
PI Fitzky B, Glossmann H, Moebius F;
XX
DR WPI; 1999-191430/17.
DR N-PSDB; X23387.
XX
PT Human Delta7-sterol reductase polypeptide - useful for diagnosis or
PT treatment of genetic defects e.g. hereditary Smith-Lemli-Opitz
PT syndrome
XX
PS Disclosure; Page 29-30; 62pp; German.
XX
CC This invention describes a human Delta7-sterol reductase. The encoding
CC DNA can be used to diagnose or correct human Delta7-sterol reductase gene
CC defects e.g. hereditary Smith-Lemli-Opitz syndrome, and to produce the
CC recombinant Delta7-sterol reductase polypeptide, which can be used to
CC replace a defective Delta7-sterol reductase enzyme in humans or other
CC animals. It is also useful to screen for Delta7-sterol reductase

```

XX PS Example 1; Page 18-21; 25pp; English.
XX
XX CC This invention describes a novel cold active protease, CP70 from
CC Flavobacterium balustinum. The product of the invention is useful
CC in detergent compositions for use in cold temperature wash water.
CC The characterization of the polynucleotide facilitates the production
CC of an alternative relatively large source of cold active protease.
XX
XX SQ Sequence 699 AA;

Query Match 71.4%; Score 30; DB 20; Length 699;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SGVTSQRT 8
Db 482 sgvtstkt 489
||||| |

RESULT 7
Y02653
ID Y02653 standard; Protein; 265 AA.
XX
AC Y02653;
XX
DT 11-JUN-1999 (first entry)
XX
DE Human secreted protein encoded by gene 4 clone HCHAA63.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX
PN WO9902546-A1.
XX
PD 21-JAN-1999.
XX
PF 07-JUL-1998; 98WO-US13684.
XX
PR 12-SEP-1997; 97US-0058785.
PR 08-JUL-1997; 97US-0051916.
PR 08-JUL-1997; 97US-0051918.
PR 08-JUL-1997; 97US-0051919.
PR 08-JUL-1997; 97US-0051920.
PR 08-JUL-1997; 97US-0051925.
PR 08-JUL-1997; 97US-0051926.
PR 08-JUL-1997; 97US-0051928.
PR 08-JUL-1997; 97US-0051929.
PR 08-JUL-1997; 97US-0051930.
PR 08-JUL-1997; 97US-0051931.
PR 08-JUL-1997; 97US-0051932.
PR 08-JUL-1997; 97US-0052732.
PR 08-JUL-1997; 97US-0052733.
PR 08-JUL-1997; 97US-0052793.
PR 08-JUL-1997; 97US-0052795.
PR 08-JUL-1997; 97US-0052803.
PR 18-AUG-1997; 97US-0055684.
PR 18-AUG-1997; 97US-0055722.
PR 18-AUG-1997; 97US-0055723.
PR 18-AUG-1997; 97US-0055947.
PR 18-AUG-1997; 97US-0055948.
PR 18-AUG-1997; 97US-0055949.
PR 18-AUG-1997; 97US-0055950.
PR 18-AUG-1997; 97US-0055953.

```

```

PR 18-AUG-1997; 97US-0055954.
PR 18-AUG-1997; 97US-0055964.
PR 18-AUG-1997; 97US-0055984.
PR 18-AUG-1997; 97US-0056360.
PR 12-SEP-1997; 97US-0058660.
PR 12-SEP-1997; 97US-0058661.
PR 12-SEP-1997; 97US-0058664.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Brewer LA, Ebner R, Fischer CL, Kyaw H, Lafleur DW, Li Y, Moore PA;
PI Olsen HS, Rosen CA, Ruben SM, Shi Y, Soppet DR, Zeng Z;
XX
DR WPI; 1999-120770/10.
DR N-PSDB; X27314.
XX
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX
PS Claim 11; Page 320-321; 464pp; English.
XX
CC This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin FC
CC portion (e.g. X27302) for increasing the stability of the fused protein
CC as compared to the human protein only.
CC The invention relates to 123 novel genes and their fragments (nucleic
CC acid sequences: X27311-X27449; amino acid sequences Y02650-Y02788) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 123 polynucleotides, based on
CC which tissues they are most highly expressed in (see X27311 for described
CC uses).
XX
SQ Sequence 265 AA;

Query Match 69.0%; Score 29; DB 20; Length 265;
Best Local Similarity 66.7%; Pred. No. 62;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SGVTSQRTA 9
Db 115 saltgrta 123
|||||

RESULT 8
Y29333
ID Y29333 standard; Protein; 475 AA.
XX
AC Y29333;
XX
DT 29-SEP-1999 (first entry)
XX
DE Human secreted protein clone pj323_2 protein sequence.
XX
KW Human; secreted protein; nutrition; cytokine; cell proliferation;
KW differentiation; immune stimulating; vaccine; suppression; gene therapy;
KW haematopoiesis regulation; tissue growth; activin; inhibin; cadherin;
KW chemotactic; chemokinetic; haemostatic; thrombolytic; anti-inflammatory;
KW tumour invasion suppressor; tumour inhibition.
XX
OS Homo sapiens.
XX
PN WO9937674-A1.
XX
PD 29-JUL-1999.
XX
PF 21-JAN-1999; 99WO-US01404.
XX

```

DE S. pneumoniae ISL2 protein.
KW Streptococcus pneumoniae strain 0100993; vaccine; immune response;
KW streptococcal infection; pneumococcal.
XX
OS Streptococcus pneumoniae.
XX WO9737026-A1.
PN
PD 09-OCT-1997.
XX
XX 01-APR-1997; 97WO-US05306.
PF
PR 22-AUG-1996; 96US-0025788.
PR 02-APR-1996; 96US-0014690.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Black MT, Hodgson JE, Knowles DJC, Nicholas RO;
PI Stodola RK;
PI
DR WPI: 1997-503111/46.
DR N-PSDB; X30727.
XX
PT Nucleic acids encoding pneumococcal polypeptide(s) - useful in
PT vaccines, drug screening, etc
XX
PS Claim 6; Page 189-190; 354pp; English.
XX
CC X30724 to X30946 represent genomic DNA sequences isolated from
CC Streptococcus pneumoniae strain 0100993. These genomic DNA sequences
CC encode the novel proteins given in Y1114 to Y11367. The proteins,
CC isolated from Streptococcus pneumoniae, can be used in vaccines against
CC streptococcal infections and in assays for identifying compounds that
CC inhibit or activate the activity of the proteins. The antagonists can
CC be used to treat an individual having need to inhibit a bacterial
CC protein. Vectors expressing the proteins can be used to induce a
CC protective immune response in mammals.
XX
XX Sequence 132 AA;
SQ
Query Match 71.4%; Score 30; DB 18; Length 132;
Best Local Similarity 87.5%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 SGVTOSRT 8
Db ||||| |||
96 sgvtisrt 103
RESULT 5
Y8585
ID Y85855 standard; Protein; 132 AA.
AC Y85855;
XX
XX
DT 10-APR-2000 (first entry)
XX
DE S. pneumoniae derived protein #64.
XX
KW Treatment; prevention; disease; diagnosis; gene therapy; screening;
KW bacterial; antimicrobial; antibiotic; pathogenesis; infection.
XX
OS Streptococcus pneumoniae.
XX
XX WO9806734-A1.
PN
XX
PD 19-FEB-1998.
XX
XX 15-AUG-1997; 97WO-US14436.
PF
XX
PR 16-AUG-1996; 96US-0024022.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;
PI Stodola RK;
XX
XX WPI: 1998-159452/14.
DR N-PSDB; Z96232.
XX
PT Streptococcus pneumoniae proteins and related DNA - useful for
PT screening compounds for antibacterial activity
XX
PS Claim 5; Page 362-363; 640pp; English.
XX
CC This invention describes novel isolated Streptococcus pneumoniae
CC polynucleotides (see Z96173-Z96494) and their encoded proteins (see
CC Y85792-Y86182). The DNA, vectors and host cells described in the method
CC of the invention are useful for the recombinant expression of the
CC polypeptides. The polypeptides are useful for treatment or prevention of
CC disease, or diagnosis of disease related to expression or activity of
CC such a polypeptide. They can also be used to screen for compounds which
CC interact with and inhibit or activate such a polypeptide. The
CC polypeptides (or DNA encoding them, via gene therapy) are also useful
CC for inducing an immunological response in a mammal. The antagonists are
CC useful to inhibit such bacterial polypeptides. The polypeptides are
CC particularly useful to identify antimicrobial compounds and antibiotics.
CC They are also useful to determine their role in pathogenesis of
CC infection, dysfunction and disease.
XX
XX Sequence 132 AA;
SQ
Query Match 71.4%; Score 30; DB 19; Length 132;
Best Local Similarity 87.5%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 SGVTOSRT 8
Db ||||| |||
96 sgvtisrt 103
RESULT 6
Y08471
ID Y08471 standard; Protein; 699 AA.
XX
AC Y08471;
XX
DT 28-JUL-1999 (first entry)
XX
DE F. balustinum CP70 protein.
XX
KW CP70; cold active protease; detergent; cold temperature wash water.
XX
OS Flavobacterium balustinum.
XX
XX WO9925848-A1.
PN
XX
PD 27-MAY-1999.
XX
PF 14-NOV-1997; 97WO-US20843.
XX
PR 14-NOV-1997; 97WO-US20843.
XX
PA (PROC) PROCTER & GAMBLE CO.
XX
PI Kitado H, Yoshikawa A;
XX
DR WPI: 1999-347483/29.
DR N-PSDB; V72330.
XX
XX Newly isolated CP70 polynucleotide exhibiting characteristics of
PT Deposit No. FERM BP-6154

PT hyper-calcaemia, hypo- or hyper-phosphatemia, kidney stones, etc
 XX Claim 9; Fig 1A-E; 62pp; English.
 XX
 CC A novel 7-transmembrane receptor (W12695) has been identified as a
 CC human G-protein parathyroid hormone (PTH) receptor, designated
 CC HL7DG74. It shows 48.2% homology to the human PTH receptor. Its
 CC amino acid sequence was deduced from a cDNA clone (T59619) isolated
 CC from a human T cell lymphoma tissue cDNA library. Recombinant
 CC HL7DG74 can be produced in transformed host cells and used to
 CC screen for (ant)agonist cpts. Agonists may be used to prevent or
 CC treat e.g. hypocalcaemia, hyperphosphatemia, hypoparathyroidism
 CC and chronic tetany by stimulating an increase in serum calcium
 CC levels. Antagonists can be used to inhibit the receptor e.g. for
 CC the treatment of osteoporosis, hypercalcaemia, hypoparathyroidism,
 CC hypophosphatemia, kidney stone, nephrolithiasis.
 XX
 SQ Sequence 541 AA;

Query Match 100.0%; Score 42; DB 18; Length 541;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SGVTSRTA 9
 |||||
 Db 496 sgvtgsrta 504

RESULT 2
 R85754
 ID R85754 standard; Protein; 885 AA.
 AC R85754;
 XX
 XX 13-FEB-1996 (first entry)
 DE Human axl receptor (alternative version).
 XX
 XX Ax1 receptor; receptor tyrosine-kinase; leukaemia.
 KW Homo sapiens.
 OS
 XX US5468634-A.
 PN
 XX 21-NOV-1995.
 PD
 XX 24-JUN-1991; 91US-0718572.
 PF
 XX 24-JUN-1991; 91US-0718572.
 PR
 XX 13-JAN-1995; 95US-0372892.
 XX
 XX (UYNC-) UNIV NORTH CAROLINA.
 PA
 XX
 XX Liu ET;
 PI
 XX WPI; 1996-010099/01.
 DR
 DR N-PSDB; T02583.
 XX
 XX DNA encoding mammalian axl receptor having tyrosine kinase activity
 PT - useful in diagnosis and treatment of tumours
 XX
 XX Disclosure; Column 29-36; 21pp; English.
 PS
 XX A cDNA clone (T02583) isolated from normal human fibroblast IMR-90
 CC cells encoded an altered version of the novel axl receptor
 CC tyrosine-kinase (R85754), probably a splice variant, which in
 CC comparison to the previously isolated axl sequence (see R85753)
 CC lacked amino acids 429-437 and had an L338K mutation.
 XX
 XX Sequence 885 AA;

Query Match 76.2%; Score 32; DB 17; Length 885;
 Best Local Similarity 87.5%; Pred. No. 48;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVTQSRTA 9
 |||||
 Db 212 gvttsrta 219

RESULT 3
 R85753
 ID R85753 standard; Protein; 894 AA.
 AC R85753;
 XX
 XX 13-FEB-1996 (first entry)
 DT
 XX Human axl receptor.
 DE
 XX Ax1 receptor; receptor tyrosine-kinase; leukaemia.
 KW Homo sapiens.
 OS
 XX US5468634-A.
 PN
 XX 21-NOV-1995.
 PD
 XX 24-JUN-1991; 91US-0718572.
 PF
 XX 24-JUN-1991; 91US-0718572.
 PR
 XX 13-JAN-1995; 95US-0372892.
 XX
 XX (UYNC-) UNIV NORTH CAROLINA.
 PA
 XX
 XX Liu ET;
 PI
 XX WPI; 1996-010099/01.
 DR
 DR N-PSDB; T02582.
 XX
 XX DNA encoding mammalian axl receptor having tyrosine kinase activity
 PT - useful in diagnosis and treatment of tumours
 XX
 XX Disclosure; Column 17-24; 21pp; English.
 PS
 XX A transforming gene in the DNA of patients with chronic myelogenous
 CC leukaemia was used to generate explant tumour cells in nude mice.
 CC Clone 1-4 (see T02582) isolated from a cDNA library of a tumour
 CC explant encoded the novel axl receptor (R85753), which has
 CC tyrosine-kinase activity involved in the progression of CML to acute
 CC phase leukaemia. Recombinant, soluble and chimeric proteins based
 CC on the axl receptor have been produced in mammalian and insect host
 CC cells.
 XX
 XX Sequence 894 AA;

Query Match 76.2%; Score 32; DB 17; Length 894;
 Best Local Similarity 87.5%; Pred. No. 48;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVTQSRTA 9
 |||||
 Db 212 gvttsrta 219

RESULT 4
 Y11118
 ID Y11118 standard; Protein; 132 AA.
 AC Y11118;
 XX
 XX 20-MAY-1999 (first entry)
 DT
 XX

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:51:50 ; Search time 138.73 Seconds
(without alignments)
2.218 Million cell updates/sec

Title: US-09-236-468A-2_COPY_496_504

Perfect score: 42

Sequence: 1 SGVTQSRSA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_36.*
1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	541	W12695	G-protein parathyr
2	32	76.2	885	17 R85754	Human axl receptor
3	32	76.2	894	17 R85753	Human axl receptor
4	30	71.4	132	18 Y11118	S. pneumoniae ISL2
5	30	71.4	132	19 Y85855	S. pneumoniae deri
6	30	71.4	699	20 Y08471	F. balustinum Cp70
7	29	69.0	265	20 Y02653	Human secreted pro
8	29	69.0	475	20 Y29333	Human secreted pro
9	29	69.0	475	20 W93572	Human Delta7-ster
10	29	69.0	475	20 W93573	Human Delta7-ster
11	28	66.7	96	12 R10175	Rap (R) protein en
12	28	66.7	96	17 R94544	Native Vpr protein

13	28	66.7	96	21 Y53247	HIV-1 LAI strain V
14	28	66.7	113	13 R26967	Human T lymphocyte
15	28	66.7	113	13 R26966	Human T lymphocyte
16	28	66.7	241	19 W55085	Streptococcus pneu
17	28	66.7	288	21 Y94633	HIF-1alpha variant
18	28	66.7	301	21 Y94634	HIF-1alpha variant
19	28	66.7	322	17 R95003	CagB antigenic pol
20	28	66.7	346	21 Y75539	Neisseria gonorrhoe
21	28	66.7	361	20 W99387	S. erythraea dTDP-4
22	28	66.7	613	21 Y94630	HIF-1alpha variant
23	28	66.7	623	19 W75773	Amino acid sequenc
24	28	66.7	642	21 Y84168	A variant of human
25	28	66.7	669	21 Y84167	A variant of human
26	28	66.7	697	21 Y84166	A variant of human
27	28	66.7	701	21 Y84173	A variant of human
28	28	66.7	710	21 Y84172	A variant of human
29	28	66.7	724	21 Y84171	A variant of human
30	28	66.7	749	21 Y84170	A variant of human
31	28	66.7	756	21 Y94635	HIF-1alpha variant
32	28	66.7	789	21 Y84169	A variant of human
33	28	66.7	805	18 W06558	Hypoxia inducible
34	28	66.7	813	21 Y94636	HIF-1alpha variant
35	28	66.7	826	18 W06557	Human hypoxia indu
36	28	66.7	826	19 W80418	Amino acid sequenc
37	28	66.7	826	20 Y06289	Human transcrip
38	28	66.7	826	21 Y94640	Human hypoxia-indu
39	28	66.7	826	21 Y69407	A wild type human
40	28	66.7	943	19 W78900	Rat UNC-5 homology
41	28	66.7	1141	14 R31961	Human cardiac cG
42	27	64.3	32	18 W08786	HVRI region of HCV
43	27	64.3	92	16 R79171	TCR Vbeta5.3. Hom
44	27	64.3	205	20 W92525	Human transcrip
45	27	64.3	206	19 W60957	Streptococcus pneu

ALIGNMENTS

RESULT 1

W12695
ID W12695 standard; Protein; 541 AA.

XX
AC W12695;

XX
DT 31-MAY-1997 (first entry)

XX
DE G-protein parathyroid hormone receptor HLTG74.

XX
KW G-protein parathyroid hormone receptor; HLTG74; parathormone; PTH;

XX
KW calcium; signal transduction; agonist; antagonist; hypocalcaemia;

XX
KW hyperphosphataemia; hypoparathyroidism; chronic tetany;

XX
KW osteoporosis; hypercalcaemia; hypoparathyroidism; hypophosphataemia;

XX
KW kidney stone; nephroliasis; therapy; diagnosis.

XX
OS Homo sapiens.

XX
PN W06639433-AA.

XX
PD 12-DEC-1996.

XX
PF 05-JUN-1995; 95WO-US07085.

XX
PR 05-JUN-1995; 95WO-US07085.

XX
PA (HUMA-) HUMAN GENOME SCI INC.

XX
PI Li Y, Rosen CA, Ruben SM, Soppet DR;

XX
WPI: 1997-043068/04.

XX
DR N-PSDB; T59619.

PT Human G-protein parathyroid hormone receptor, HLTG74 - used to identify (ant)agonists, used in the treatment of hypo- or

Db 706 SKIPRSPD 714

Search completed: November 8, 2000, 08:56:20
Job time: 531 sec

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=GLIOBLASTOMA;
RA Owens R.J., Lumb S.M., Rees-Milton K., Russell A., Baldock D.,
RA Lang V., Crabbe T., Ballesteros M., Perry M.J.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U88713; AAC51916.1; -.
DR EMBL; U88712; AAC51915.1; -.
DR INTERPRO: IPR002073; -.
DR PFAM: PF00233; PDEase; 1.
DR PROSITE; PS00126; PDEASE_I; 1.
FT NON_TER 1
SQ SEQUENCE 700 AA; 77600 MW; B39732360C2E18A5 CRC64;

Query Match 64.4%; Score 38; DB 4; Length 700;
Best Local Similarity 66.7%; Pred. No. 49;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
Db 624 SKIPRSPD 632
:|:||||:|

RESULT 13
Q9UPJ6 PRELIMINARY; PRT; 710 AA.
AC Q9UPJ6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE CNAC_HUMAN (EC 3.1.4.17).
CN PDE4C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Frankhelm M., Amico-Keller G.,
RA Cosfield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Kronmiller B., Arellano A., Saunders C., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RA "Sequence analysis of an ~600 kb region in 19p13.1 between JAK3 and
RT PDE4C.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005759; AAC3047.1; -.
DR INTERPRO: IPR002073; -.
DR PFAM: PF00233; PDEase; 1.
DR PRINTS; PR00387; PDIESTERASE1.
DR PROSITE; PS00126; PDEASE_I; 1.
KW Hydrolase.
SQ SEQUENCE 710 AA; 79699 MW; D568395BBC19A2FA CRC64;

Query Match 64.4%; Score 38; DB 4; Length 710;
Best Local Similarity 66.7%; Pred. No. 50;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
Db 634 SKIPRSPD 642
:|:||||:|

RESULT 14
Q9UN46 PRELIMINARY; PRT; 712 AA.
ID Q9UN46
AC Q9UN46;

```

```

DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE CAMP SPECIFIC PHOSPHODIESTERASE PDE4C1 VARIANT.
CN PDE4C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA Sullivan M., Olsen A.S., Houslay M.D.;
RT "Genomic organization of the human cyclic AMP specific
RT phosphodiesterase PDE4C gene and its chromosomal localization to
RT 19p13.1, between the genes for RAB3A and JUND.";
RL Cell. Signal. 0:0-0(1999).
DR EMBL; AF157816; AAD47053.1; -.
DR EMBL; AF157811; AAD47053.1; JOINED.
DR EMBL; AF157814; AAD47053.1; JOINED.
DR EMBL; AF157815; AAD47053.1; JOINED.
DR INTERPRO: IPR002073; -.
DR PFAM: PF00233; PDEase; 1.
DR PRINTS; PR00387; PDIESTERASE1.
DR PROSITE; PS00126; PDEASE_I; 1.
SQ SEQUENCE 712 AA; 79935 MW; 2932116C9D70B655 CRC64;

Query Match 64.4%; Score 38; DB 4; Length 712;
Best Local Similarity 66.7%; Pred. No. 50;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
Db 636 SKIPRSPD 644
:|:||||:|

RESULT 15
Q76105 PRELIMINARY; PRT; 782 AA.
ID Q76105;
AC Q76105;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE PDB4C-791 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Frankhelm M., Amico-Keller G.,
RA Cosfield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Kronmiller B., Arellano A., Saunders C., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RA "Sequence analysis of an ~600 kb region in 19p13.1 between JAK3 and
RT PDE4C.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005759; AAC83050.1; -.
DR INTERPRO: IPR002073; -.
DR PFAM: PF00233; PDEase; 1.
DR PROSITE; PS00126; PDEASE_I; 1.
FT NON_TER 1
SQ SEQUENCE 782 AA; 87342 MW; 42B7F735F1C17D24 CRC64;

Query Match 64.4%; Score 38; DB 4; Length 782;
Best Local Similarity 66.7%; Pred. No. 55;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
Db 636 SKIPRSPD 644
:|:||||:|

```

```
Best Local Similarity 70.0%; Pred. No. 35;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KLPRSPADSL 11
    |||||:::|
Db 286 KLPRSPREAL 295

RESULT 9
O43850 PRELIMINARY; PRT; 518 AA.
AC O43850;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE PDE4C-DELTA54 (FRAGMENT).
GN PDE4C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE: 98007880.
RA Oboernolte R., Ratzliff J., Baecker P.A., Daniels D.V., Zuppan P.,
RA Jarnagin K., Shelton E.R.;
RT "Multiple splice variants of phosphodiesterase PDE4C cloned from human
RT lung and testis.";
RL Biochim. Biophys. Acta 1353:287-297(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Frankheim M., Amico-Keller G.,
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Kronmiller B., Arellano A., Saunders C., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RA "Sequence analysis of an ~600 kb region in 19p13.1 between JAK3 and
RT PDE4C.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; U66348; AAB96877.1; -.
DR EMBL; AC005759; AAC83051.1; -.
DR INTERPRO; IPR002073; -.
DR PFAM; PF00233; PDEase; 1.
DR PROSITE; PS00126; PDEASE_I; 1.
FT NON_TER
SQ SEQUENCE 518 AA; 58116 MW; DDC5BE289C1192C CRC64;

Query Match 64.4%; Score 38; DB 4; Length 518;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
    :|||::|
Db 442 SKIPRSPD 450

RESULT 10
O43850 PRELIMINARY; PRT; 606 AA.
AC O43850;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE CAMP SPECIFIC PHOSPHODIESTERASE PDE4C2 VARIANT.
GN PDE4C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
```

```
SEQUENCE FROM N.A.
RA Sullivan M., Olsen A.S., Houslay M.D.;
RT "Genomic organization of the human cyclic AMP specific
RT phosphodiesterase PDE4C gene and its chromosomal localization to
RT 19p13.1, between the genes for RAB3A and JUND.";
RL Cell. Signal. 0:0-0(1999).
DR EMBL; AF157816; AAD47055.1; -.
DR EMBL; AF157814; AAD47055.1; JOINED.
DR EMBL; AF157815; AAD47055.1; JOINED.
DR INTERPRO; IPR002073; -.
DR PFAM; PF00233; PDEase; 1.
DR PRINTS; PR00387; PDIESTERASE1.
DR PROSITE; PS00126; PDEASE_I; 1.
SQ SEQUENCE 606 AA; 67824 MW; OC42FC9F22E3B798 CRC64;

Query Match 64.4%; Score 38; DB 4; Length 606;
Best Local Similarity 66.7%; Pred. No. 43;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
    :|||::|
Db 530 SKIPRSPD 538

RESULT 11
O43850 PRELIMINARY; PRT; 680 AA.
AC O43850;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE CAMP SPECIFIC PHOSPHODIESTERASE PDE4C3 VARIANT.
GN PDE4C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA Sullivan M., Olsen A.S., Houslay M.D.;
RT "Genomic organization of the human cyclic AMP specific
RT phosphodiesterase PDE4C gene and its chromosomal localization to
RT 19p13.1, between the genes for RAB3A and JUND.";
RL Cell. Signal. 0:0-0(1999).
DR EMBL; AF157816; AAD47054.1; -.
DR EMBL; AF157812; AAD47054.1; JOINED.
DR EMBL; AF157814; AAD47054.1; JOINED.
DR EMBL; AF157815; AAD47054.1; JOINED.
DR INTERPRO; IPR002073; -.
DR PFAM; PF00233; PDEase; 1.
DR PRINTS; PR00387; PDIESTERASE1.
DR PROSITE; PS00126; PDEASE_I; 1.
SQ SEQUENCE 680 AA; 75647 MW; D555883F3341FFID CRC64;

Query Match 64.4%; Score 38; DB 4; Length 680;
Best Local Similarity 66.7%; Pred. No. 48;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
    :|||::|
Db 604 SKIPRSPD 612

RESULT 12
P78505 PRELIMINARY; PRT; 700 AA.
ID P78505
AC P78505;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE PHOSPHODIESTERASE 4C (FRAGMENT).
OS Homo sapiens (Human).
```

```
AC Q9UPJ5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE PDE4C-426.
GN PDE4C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Trankheim M., Amico-Keller G.,
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Krommiller B., Arellano A., Saunders C., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RA "Sequence analysis of an ~600 kb region in 19p13.1 between JAK3 and
RT PDE4C.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005759; AAC83049.1; -.
DR INTERPRO; IPR002073; -.
DR PFAM; PF00233; PDEase; 1.
DR PRINTS; PR00387; PDIESTERASE1.
DR PROSITE; PS00126; PDEASE_I; 1.
SQ SEQUENCE 426 AA; 47870 MW; D29CEA570AC86AA9 CRC64;

Query Match 64.4%; Score 38; DB 4; Length 426;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLPRSPAD 9
Db 350 SKIPRSPD 358
:|:||||:|

RESULT 6
ID O43851 PRELIMINARY; PRT; 427 AA.
AC O43851;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE PDE4C-DELTA109 (FRAGMENT).
GN PDE4C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA Obernolte R., Ratcliff J., Baecker P.A., Daniels D.V., Zuppan P.,
RA Jarnagin K., Shelton E.R.;
RT "Multiple splice variants of phosphodiesterase PDE4C cloned from human
RT lung and testis.";
RL Biochim. Biophys. Acta 1353:287-297(1997).
DR EMBL; U66349; AAB96878.1; -.
DR INTERPRO; IPR002073; -.
DR PFAM; PF00233; PDEase; 1.
DR PROSITE; PS00126; PDEASE_I; 1.
FT NON_TER 1
SQ SEQUENCE 427 AA; 48217 MW; 131D4129F5362FFB CRC64;

Query Match 64.4%; Score 38; DB 4; Length 427;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLPRSPAD 9
```

```
Db 351 SKIPRSPD 359
:|:||||:|

RESULT 7
ID O76104 PRELIMINARY; PRT; 427 AA.
AC O76104;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE PDE4C-DELTA109.
GN PDE4C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Trankheim M., Amico-Keller G.,
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Krommiller B., Arellano A., Saunders C., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RA "Sequence analysis of an ~600 kb region in 19p13.1 between JAK3 and
RT PDE4C.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005759; AAC83048.1; -.
DR INTERPRO; IPR002073; -.
DR PFAM; PF00233; PDEase; 1.
DR PROSITE; PS00126; PDEASE_I; 1.
SQ SEQUENCE 427 AA; 48245 MW; 120391E6P5F8312B CRC64;

Query Match 64.4%; Score 38; DB 4; Length 427;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLPRSPAD 9
Db 351 SKIPRSPD 359
:|:||||:|

RESULT 8
ID O75502 PRELIMINARY; PRT; 496 AA.
AC O75502;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE SARCO-/ENDOPLASMIC RETICULUM CA-ATPASE 3 (FRAGMENT).
GN ATP2A3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA Poch E., Leach S., Snape S., Cacic T., Macleannan D.H., Lytton J.;
RT "Functional characterization of alternatively spliced human SERCA3
RT transcripts.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF068221; AAC24526.1; -.
DR INTERPRO; IPR001757; -.
DR PFAM; PF00122; E1-E2_ATPase; 1.
FT NON_TER 1
FT VARIANT 284 M -> I.
SQ SEQUENCE 496 AA; 54633 MW; 4BB41F7146DA3DFE CRC64;

Query Match 64.4%; Score 38; DB 4; Length 496;
```

```

RESULT 2
Q9VUX6 PRELIMINARY: PRT: 1023 AA.
AC Q9VUX6;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
DE CG5891 PROTEIN.
OS Drosophila melanogaster (Fruit fly).
NC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE; 20196006.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballev R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003528; AAF49547.1;
DR FLYBASE; FBgn0036561; CG5891.
DR INTERPRO; IPR002110;
DR PFAM; PF00023; ank; 5.
SQ SEQUENCE 1023 AA; 109468 MW; E5CD23078274BB9A CRC64;

Query Match 71.2%; Score 42; DB 5; Length 1023;
Best Local Similarity 72.7%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KLPRSPADSLT 12
Db 428 KLPRTPSDSPT 438

RESULT 3
Q9WPG8 PRELIMINARY: PRT: 596 AA.
AC Q9WPG8;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE DNA-DEPENDENT DNA POLYMERASE (FRAGMENT).
OS Columbiv herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC unclassified Herpesviridae.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KP 21/23;
RX MEDLINE; 9383789.
RA Ehlers B., Borchers K., Grund C., Froelich K., Ludwig H., Buhk H.-J.;
RA "Detection of new DNA polymerase genes of known and potentially novel
RA herpesviruses by PCR with degenerate and deoxynosine-substituted
RA primers.";
RL Virus Genes 18:211-220(1999).
DR EMBL; AF141890; AAD30145.1;
DR INTERPRO; IPR002064;
DR PFAM; PF00136; DNA_POL_B; 1.
DR PRINTS; PR00106; DNAPOLB.
FT NON_TER 1
FT NON_TER 596
SQ SEQUENCE 596 AA; 16178558DBB86449 CRC64;

Query Match 66.1%; Score 39; DB 12; Length 596;
Best Local Similarity 77.8%; Pred. No. 27;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
Db 503 AKIPRSPED 511

RESULT 4
Q9Z258 PRELIMINARY: PRT: 1237 AA.
AC Q9Z258;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
DE POTASSIUM CHANNEL SUBUNIT.
GN SLACK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RA Joiner W.J., Tang M.D., Wang L.-Y., Dworetzky S.I., Bolissard C.G.,
RA Gan L., Gribkoff V.K., Kaczmarek L.K.;
RA "Formation of intermediate-conductance calcium-activated potassium
RA channels by interaction of Slack and Slo subunits.";
RL Nat. Neurosci. 0:0-0(1999).
DR EMBL; AF089730; AAC8350.1;
DR INTERPRO; IPR001622;
KW Ionic channel.
SQ SEQUENCE 1237 AA; 139614 MW; E1B4A7A77EB612B CRC64;

Query Match 66.1%; Score 39; DB 11; Length 1237;
Best Local Similarity 77.8%; Pred. No. 56;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
Db 4 AKLPRSPSE 12

RESULT 5
Q9UPJ5 PRELIMINARY: PRT: 426 AA.
ID Q9UPJ5

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:56:18 ; Search time 152.43 Seconds
(without alignments)
7.351 Million cell updates/sec

Title: US-09-236-468A-2_COPY_476_487

Perfect score: 59

Sequence: 1 AKLPRSPADSLT 12

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 297973

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_14:*

- 1: sp-archaea:*
- 2: sp-bacteria:*
- 3: sp-fungi:*
- 4: sp-human:*
- 5: sp-invertebrate:*
- 6: sp-mammal:*
- 7: sp-mhmc:*
- 8: sp-organelle:*
- 9: sp-phage:*
- 10: sp-plant:*
- 11: sp-rodent:*
- 12: sp-virus:*
- 13: sp-vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	42	71.2	413	2 Q08859	Q08859 Klebsiella
2	42	71.2	1023	5 Q9VUX6	Q9vux6 drosophila
3	39	66.1	596	12 Q9WPC8	Q9wpc8 columbid he
4	39	66.1	1237	11 Q9Z258	Q9z258 rattus norv
5	38	64.4	426	4 Q9UPJ5	Q9upj5 homo sapien
6	38	64.4	427	4 Q43851	Q43851 homo sapien
7	38	64.4	427	4 Q76104	Q76104 homo sapien
8	38	64.4	496	4 Q75502	Q75502 homo sapien
9	38	64.4	518	4 Q43850	Q43850 homo sapien
10	38	64.4	606	4 Q9UN44	Q9un44 homo sapien
11	38	64.4	680	4 Q9UN45	Q9un45 homo sapien
12	38	64.4	700	4 P78505	P78505 homo sapien
13	38	64.4	710	4 Q9UPJ6	Q9upj6 homo sapien
14	38	64.4	712	4 Q9UN46	Q9un46 homo sapien
15	38	64.4	782	4 Q76105	Q76105 homo sapien
16	38	64.4	791	4 Q43849	Q43849 homo sapien
17	38	64.4	998	4 Q75501	Q75501 homo sapien
18	38	64.4	1029	4 Q60900	Q60900 homo sapien
19	38	64.4	1407	2 Q69826	Q69826 streptomyce

20 37 62.7 217 2 Q9ZAI5
21 37 62.7 492 10 Q9ZW77
22 36 61.0 305 10 Q9SNV5
23 36 61.0 338 12 Q11878
24 36 61.0 339 5 Q44085
25 36 61.0 508 13 Q9YHV9
26 36 61.0 508 13 Q9PVK0
27 36 61.0 789 12 Q88167
28 36 61.0 995 4 Q00291
29 36 61.0 1744 5 Q17008
30 36 61.0 1879 5 Q19161
31 35 59.3 66 2 Q52791
32 35 59.3 66 2 Q57218
33 35 59.3 135 6 Q77696
34 35 59.3 157 4 Q16192
35 35 59.3 209 2 Q86679
36 35 59.3 238 2 Q52806
37 35 59.3 243 2 Q52817
38 35 59.3 281 2 Q9RPF6
39 35 59.3 282 2 Q87610
40 35 59.3 282 2 Q9RPB3
41 35 59.3 284 2 Q9RUZ0
42 35 59.3 284 2 Q9RC37
43 35 59.3 344 5 Q9VJP7
44 35 59.3 349 4 Q9UF43
45 35 59.3 356 3 Q74361

ALIGNMENTS

RESULT 1
Q08859 PRELIMINARY; PRT: 413 AA.
ID Q08859;
AC Q08859;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE FIMK PROTEIN.
GN FIMK.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89155420.
RA Gerlach G.F., Clegg S., Allen B.L.;
RT "Identification and characterization of the genes encoding the type 3
and type 1 fimbrial adhesins of Klebsiella pneumoniae.";
RL J. Bacteriol. 171:1262-1270(1989).
RN [2]
RP SEQUENCE FROM N.A.
RA Nicholas N.N., Clegg S.;
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; L23111; AAA25064.1; -.
DR INTERPRO; IPR000792; -.
DR INTERPRO; IPR001633; -.
DR INTERPRO; IPR002106; -.
DR PFAM; PF00563; DUF2; 1.
KW Fimbria.

SQ SEQUENCE 413 AA; 46940 MW; A29890E2A4C60655 CRC64;

Query Match 71.2%; Score 42; DB 2; Length 413;
Best Local Similarity 80.0%; Pred. No. 5.3;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LPRSPADSLT 12

|||||:|

Db 147 LPRSPANALT 156

THIS PAGE BLANK (USPTO)

```
DR PFAM; PF01977; UPF0096; 1.
KW Hypothetical protein.
SQ SEQUENCE 481 AA; 54266 MW; 6A1B9087D24AB9FB CRC64;

Query Match
Best Local Similarity 57.6%; Score 34; DB 1; Length 481;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LPRSPADSL 11
Db 417 LPPSPDLSL 425
II III III

RESULT 14
Y039 HUMAN
ID Y039 HUMAN STANDARD; PRT; 491 AA.
AC Q15054;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE HYPOTHETICAL PROTEIN KIAA0039 (FRAGMENT).
GN KIAA0039
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BONE MARROW;
RX MEDLINE; 96051387.
RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
RA Sato S., Nagase T., Seki N., Ishikawa K.-I., Tabata S.;
RT "Prediction of the coding sequences of unidentified human genes. I.
RT The coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by
RT analysis of randomly sampled cDNA clones from human immature myeloid
RT cell line KG-1.";
RL DNA Res. 1:27-35(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D26018; BAA05039.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 491 AA; 54079 MW; AD6B81F72F467CBA CRC64;

Query Match
Best Local Similarity 57.6%; Score 34; DB 1; Length 491;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AKLPRSPADS 10
Db 164 AAVPRAPAES 173
I :II:II:I

RESULT 15
YM05 YEAST
ID YM05 YEAST STANDARD; PRT; 501 AA.
AC Q04472;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 58.0 KDA PROTEIN IN ILV2-ADEL7 INTERGENIC REGION.
GN YMR115W OR YW9718.14.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomyces.

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Hunt S., Bowman S., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: TO YEAST YKL133C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z49702; CAA89752.1; -.
DR SGD; S0004721; YMR115W.
KW Hypothetical protein.
SQ SEQUENCE 501 AA; 58003 MW; 994613B288C8DB7C CRC64;

Query Match
Best Local Similarity 57.6%; Score 34; DB 1; Length 501;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AKLPRSPADSLT 12
Db 51 APIPESPANST 62
I :I III:I

Search completed: November 8, 2000, 09:03:56
Job time: 866 sec
```



```
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Wilkinson J.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: SOME, TO YEAST VKL125W AND S. POMBE SPAC18G6.10C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z35597; CAA84645.1; -.
DR WORMPEP; C36E8.1; CE00909.
KW Hypothetical protein.
SQ SEQUENCE 654 AA; 75109 MW; 36BAD7C74DD6ADE1 CRC64;

Query Match 59.3%; Score 35; DB 1; Length 654;
Best Local Similarity 55.6%; Pred. No. 69;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 KLPRSPADS 10
DB 401 RMPRAPADA 409
:::|::|::|

RESULT 12
ATC3_RAT STANDARD; PRT; 999 AA.
ID Y209_ARCFU
AC O30030;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SARCOPLASMIC/ENDOPLASMIC RETICULUM CALCIUM ATPASE ISOFORM 3
DE (EC 3.6.1.38) (CALCIUM PUMP 3) (SERCA3) (SR CA(2+)-ATPASE 3).
GN ATP2A3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE; 90036958.
RX Burk S.E., Lytton J., McLennan D.H., Shull G.E.;
RT "cDNA cloning, functional expression, and mRNA tissue distribution of
RT a third organellar Ca2+ pump."
RL J. Biol. Chem. 264:18561-18568(1989).
CC -1- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS
CC OF ATP COUPLED WITH THE TRANSPORT OF THE CALCIUM. TRANSPORTS
CC CALCIUM IONS FROM THE CYTOSOL INTO THE SARCOPLASMIC/ENDOPLASMIC
CC RETICULUM. HAS A CENTRAL ROLE IN INTRACELLULAR CALCIUM SIGNALING.
CC -1- CATALYTIC ACTIVITY: ATP + H2O -> ADP + ORTHOPHOSPHATE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
CC (E1-E2 ATPASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M30581; AAA42131.1; -.
DR PIR; A34307; A34307.
DR INTERPRO; IPR001757; -.
DR PFAM; PF00122; E1-E2 ATPase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PROSITE; PS00154; ATPASE_E1_E2; 1.
KW Hydrolase; Calcium transport; Transmembrane; Phosphorylation;
```

```
KW Magnesium; ATP-binding.
FT DOMAIN 1 59
FT TRANSMEM 60 78
FT DOMAIN 79 86
FT TRANSMEM 87 107
FT TRANSMEM 108 259
FT TRANSMEM 298 315
FT TRANSMEM 316 772
FT TRANSMEM 773 795
FT TRANSMEM 796 836
FT TRANSMEM 837 859
FT TRANSMEM 860 894
FT TRANSMEM 895 913
FT TRANSMEM 914 930
FT TRANSMEM 931 950
FT TRANSMEM 951 999
FT MOD_RES 351 351
FT BINDING 515 515
SQ SEQUENCE 999 AA; 109359 MW; 8A6AEE29021AA379 CRC64;

Query Match 59.3%; Score 35; DB 1; Length 999;
Best Local Similarity 60.0%; Pred. No. 11e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 KLPRSPADSL 11
DB 819 KLPRNPREAL 828
||||:|::|

RESULT 13
Y209_ARCFU STANDARD; PRT; 481 AA.
ID Y209_ARCFU
AC O30030;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL PROTEIN AF0209.
GN AF0209.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE; 98049343.
RA Klenk H.P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uitterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- SIMILARITY: BELONGS TO THE UPF0096 (VDCC) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001091; AAB91023.1; -.
DR TIGR; AF0209; -.
DR INTERPRO; IPR002830; -.
```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC -----
DR EMBL; M16964; AAA98283.1; ALT_INIT.
KW Hypothetical protein; Periplasmic; Signal; Plasmid.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 329 HYPOTHETICAL 34.5 KDA PROTEIN IN
FT CLCB-CLCD INTERGENIC REGION.
SQ SEQUENCE 329 AA; 34504 MW; 819DCADA915683BE CRC64;

Query Match 59.3%; Score 35; DB 1; Length 329;
Best Local Similarity 58.3%; Pred. No. 33;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AKLPRSPADSLT 12
Db 81 AEVARSADGYT 92
I::|::|::|::|

RESULT 9
CG12_HUMAN STANDARD; PRT; 331 AA.
ID O95236;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE TNF-INDUCIBLE PROTEIN CG12_1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-ENDOTHELIAL CELLS;
RX MEDLINE; 99252096.
RA Horrevoets A.J.G., Fontijn R.D., van Zonneveld A.J., de Vries C.J.M.,
RA ten Cate J.W., Pannekoek H.;
RT "Vascular endothelial genes that are responsive to tumor necrosis
RT factor-alpha in vitro are expressed in atherosclerotic lesions,
RT including inhibitor of apoptosis protein-1, stannin, and two novel
RT genes";
RL Blood 93:3418-3431(1999).
CC -!- TISSUE SPECIFICITY: ENDOTHELIAL-CELLS SPECIFIC.
CC -!- INDUCTION: IN VITRO, IS RESPONSIVE TO TUMOR NECROSIS FACTOR.
CC -!- SIMILARITY: TO APOLIPOPROTEIN L.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

DR EMBL; AF070675; AACB3233.1; --
SQ SEQUENCE 331 AA; 36545 MW; C7703BAE5D0302FA CRC64;

Query Match 59.3%; Score 35; DB 1; Length 331;
Best Local Similarity 63.6%; Pred. No. 33;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AKLPRSPADSL 11
Db 42 AELPRDEADAL 52
I::|::|::|::|

RESULT 10
APL_HUMAN STANDARD; PRT; 383 AA.
ID APL_HUMAN
AC O14791; O60804;
DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE APOLIPOPROTEIN L PRECURSOR (APO-L).
GN APOL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 13-48.
RC TISSUE-PANCREAS;
RX MEDLINE; 97467346.
RA Duchateau P.N., Pullinger C.R., Orellana R.E., Kunitake S.T.,
RA Naya-Vigne J., O'Connor P.M., Malloy M.J., Kane J.P.;
RT "Apolipoprotein L, a new human high density lipoprotein
RT apolipoprotein expressed by the pancreas. Identification, cloning,
RT characterization, and plasma distribution of apolipoprotein L.";
RL J. Biol. Chem. 272:25576-25582(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Lloyd D.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- TISSUE SPECIFICITY: PLASMA. FOUND ON APOA-I-CONTAINING HIGH
CC DENSITY LIPOPROTEIN (HDL3). ONLY EXPRESSED IN PANCREAS.
CC -!- SIMILARITY: TO TNF-INDUCIBLE PROTEIN CG12_1.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

DR EMBL; AF019225; AAB81218.1; --
DR EMBL; Z82215; CAB05104.1; --
DR MIN; 603743; --
KW Plasma; Lipid transport; Glycoprotein; Signal; HDL.
FT SIGNAL 1 12
FT CHAIN 13 383 APOLIPOPROTEIN L.
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 9 9 R -> G (IN REF. 2).
FT CONFLICT 135 135 K -> E (IN REF. 2).
FT CONFLICT 140 140 T -> I (IN REF. 2).
FT CONFLICT 213 213 I -> M (IN REF. 2).
FT CONFLICT 240 240 K -> R (IN REF. 2).
FT CONFLICT 331 331 A -> V (IN REF. 2).
SQ SEQUENCE 383 AA; 42343 MW; 8EFF068ECOFEA462 CRC64;

Query Match 59.3%; Score 35; DB 1; Length 383;
Best Local Similarity 63.6%; Pred. No. 39;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AKLPRSPADSL 11
Db 86 AELPRNEADEL 96
I::|::|::|::|

RESULT 11
YPL_CAEEL STANDARD; PRT; 654 AA.
ID YPL_CAEEL
AC P48322;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE HYPOTHETICAL 75.1 KDA PROTEIN C36E8.1 IN CHROMOSOME III.
GN C36E8.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]

RX MEDLINE; 91108372.
 RA Vincent J.R., Ueng P.P., Lister R.M., Larkins B.A.;
 RT "Nucleotide sequences of coat protein genes for three isolates of
 RT barley yellow dwarf virus and their relationships to other luteovirus
 RT coat protein sequences.";
 RL J. Gen. Virol. 71:2791-2799(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 92013950.
 RA Larkins B.A., Lister R.M., Vincent J.R.;
 RT "Nucleotide sequence analysis and genomic organization of the NY-RPV
 RT isolate of barley yellow dwarf virus.";
 RL J. Gen. Virol. 72:2347-2355(1991).
 CC -!- FUNCTION: THIS PROTEIN IS POSSIBLY IS A VPG-PRECURSOR FROM WHICH,
 CC AT THE ONSET OF THE RNA SYNTHESIS, THE VPG MOLECULE IS RELEASED,
 CC AS IT HAS BEEN SUGGESTED FOR COMPEA MOSAIC VIRUS.
 CC -!- SIMILARITY: BELONGS TO THE LUTEOVIRUSES VPG PROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D10206; BAA01056.1; -;
 DR EMBL; X17259; CAA35161.1; -;
 DR EMBL; L25299; AAA42870.1; -;
 DR INTERPRO: IPR001964; -;
 DR PFAM; PF01659; Luteo_Vpg; 1.
 DR PRINTS; PR00912; LVIRUSORF5.
 SQ SEQUENCE 153 AA; 17209 MW; 3839B5C502C51602 CRC64;

Query Match 59.38; Score 35; DB 1; Length 153;
 Best Local Similarity 70.08; Pred. No. 15;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LPRSPADSLT 12
 ||: || |||
 Db 113 LPQPAPSLT 122

RESULT 7
 YIFB_ECOLI
 ID YIFB_ECOLI STANDARD; PRT; 202 AA.
 AC P32105;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE HYPOTHETICAL 22.6 KDA PROTEIN IN SELA-RHSA INTERGENIC REGION.
 GN YIFB.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 CC [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-K12;
 RX MEDLINE; 93259920.
 RA Zhao S., Sandt C.H., Feulner G., Vlazny D.A., Gray J.A., Hill C.W.;
 RT "Rhs elements of Escherichia coli K-12: complex composites of shared
 RT and unique components that have different evolutionary histories.";
 RL J. Bacteriol. 175:2799-2808(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE; 94316500.
 RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
 RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
 RT region from 76.0 to 81.5 minutes.";
 RL Nucleic Acids Res. 22:2576-2586(1994).
 RN [3]

RP SEQUENCE FROM N.A.
 RC STRAIN=ECOR-32;
 RA Hill C.W., Feulner G., Brody M.S., Zhao S., Sadosky A.B.,
 RA Sandt C.H.;
 RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. HSP26 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L19044; AAC95064.1; -;
 DR EMBL; U00039; AAB18569.1; -;
 DR EMBL; AE000436; AAC76616.1; -;
 DR EMBL; U16247; AAA56754.1; -;
 DR ECOGENE; EGI1762; YIFB.
 DR INTERPRO: IPR000521; -;
 DR PFAM; PF00043; GST; 1.
 KW Hypothetical protein; Transferase.
 FT CONFLICT 195 195 A -> T (IN REF. 1).
 SQ SEQUENCE 202 AA; 22545 MW; BFEA058143049D67 CRC64;

Query Match 59.38; Score 35; DB 1; Length 202;
 Best Local Similarity 63.68; Pred. No. 20;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AKLPRSPADSL 11
 | ||| | : ||
 Db 78 AMLPRDPLESL 88

RESULT 8
 YCLC_PSEPU
 ID YCLC_PSEPU STANDARD; PRT; 329 AA.
 AC Q47100; Q30846;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HYPOTHETICAL 34.5 KDA PROTEIN IN CLCB-CLCD INTERGENIC REGION
 DE PRECURSOR.
 OS Pseudomonas putida, and Pseudomonas sp. (strain B13).
 OG Plasmid pAC27.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC Pseudomonas.
 CC [1]
 RN SEQUENCE FROM N.A.
 RC SPECIES=P. PUTIDA;
 RX MEDLINE; 87260828.
 RA Frantz B., Chakrabarty A.M.;
 RT "Organization and nucleotide sequence determination of a gene cluster
 RT involved in 3-chlorocatechol degradation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:4460-4464(1987).
 RN [2]
 RP SEQUENCE OF 161-329 FROM N.A.
 RC STRAIN=B13;
 RX MEDLINE; 97315261.
 RA Kasberg T., Seibert V., Schlomann M., Reinecke W.;
 RT "Cloning, characterization, and sequence analysis of the clcE gene
 RT encoding the maleylacetate reductase of Pseudomonas sp. strain B13.";
 RL J. Bacteriol. 179:3801-3803(1997).
 CC -!- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
 CC -!- SIMILARITY: BELONGS TO THE UPF0065 (BUG) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

```

Best Local Similarity 77.8%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
Db 7 AKRPRPAD 15

RESULT 4
Y174_HUMAN
ID Y174_HUMAN STANDARD; PRT; 364 AA.
AC P53990;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL PROTEIN KIAA0174.
GN KIAA0174.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BONE MARROW;
RX MEDLINE; 96281124.
RA Nagase T., Seki N., Ishikawa K.-I., Tanaka A., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. V.
RT The coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 3:17-24(1996).
CC -!- SIMILARITY: TO YEAST YNL265C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D79996; BAAL1491.1; -.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 364 AA; 39750 MW; 0DD3C186A52A4380 CRC64;

Query Match 61.0%; Score 36; DB 1; Length 364;
Best Local Similarity 70.0%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AKLPRSPADS 10
Db 308 AKLPSRPADN 317

RESULT 5
COGM_MOUSE
ID COGM_MOUSE STANDARD; PRT; 462 AA.
AC P34960;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MACROPHAGE METALLOELASTASE PRECURSOR (EC 3.4.24.65) (MME) (MATRIX
DE METALLOPROTEINASE-12) (MMP-12).
GN MMP12 OR MME12 OR MME.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 99-125.
RC TISSUE-MACROPHAGE;
RX MEDLINE; 92165826.
RA Shapiro S.D., Griffin G.L., Gilbert D.J., Jenkins N.A.,
RA Copeland N.G., Weigus H.G., Senior R.M., Ley T.J.;
RT "Molecular cloning, chromosomal localization, and bacterial

```

```

expression of a murine macrophage metalloelastase.";
RL J. Biol. Chem. 267:4664-4671(1992).
CC -!- FUNCTION: MAY BE INVOLVED IN TISSUE INJURY AND REMODELING. HAS
CC SIGNIFICANT ELASTOLYTIC ACTIVITY.
CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF SOLUBLE AND INSOLUBLE ELASTIN.
CC SPECIFIC CLEAVAGES ARE ALSO PRODUCED AT 14-ALA-|-LEU-15 AND 16-
CC TYR-|-LEU-17 IN THE B CHAIN OF INSULIN.
CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEASE) ALSO KNOWN AS MATRXIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M82831; AAA39526.1; -.
DR PIR: A42401; A42401.
DR HSSP; P03956; ICGL.
DR MGD; MGI:97005; MMEL.
DR INTERPRO; IPR000130; -.
DR INTERPRO; IPR000585; -.
DR INTERPRO; IPR001818; -.
DR PFAM; PF00413; Peptidase_M10; 1.
DR PFAM; PF00045; hemopexin; 4.
DR PRINTS; PR00138; MATRXIN.
DR PROSITE; PS00024; HEMOPEXIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
KW Hydrolase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;
KW Extracellular matrix; Signal.
FT SIGNAL 1 17 PROBABLE.
FT PROPEP 18 98 ACTIVATION PEPTIDE.
FT CHAIN 99 462 MACROPHAGE METALLOELASTASE.
FT DOMAIN 83 90 AUTOINHIBITOR REGION.
FT METAL 272 462 HEMOPEXIN-LIKE.
FT ACT_SITE 211 212 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 215 215 BY SIMILARITY.
FT METAL 221 221 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 21 21 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 74 74 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 310 310 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 275 462 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 462 AA; 53841 MW; BB9625906F1DBEDF CRC64;

Query Match 61.0%; Score 36; DB 1; Length 462;
Best Local Similarity 63.6%; Pred. No. 31;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KLPRSPADSLT 12
Db 302 KLPGSPATNIT 312

RESULT 6
VPG_BYDVN
ID VPG_BYDVN STANDARD; PRT; 153 AA.
AC P27579;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PUTATIVE GENOME-LINKED PROTEIN PRECURSOR (VPG) (17 KDA PROTEIN).
OS Barley yellow dwarf virus (isolate NY-RPV) (BYDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Luteoviridae;
OC Luteovirus.
RN [1]
RP SEQUENCE FROM N.A.

```

DR PROSITE; PS00126; PDEASE_I; 1.
 KW Hydrolase; cAMP; Multigene family; Alternative splicing.
 SQ SEQUENCE 712 AA; 79845 MW; 0BC70E917A93F78 CRC64;

Query Match 64.4%; Score 38; DB 1; Length 712;
 Best Local Similarity 66.7%; Pred. No. 22;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
 :|:|:|:|:
 Db 636 SKIPRSPSD 644

RESULT 2

ID ATC3_HUMAN STANDARD; PRT; 999 AA.
 AC Q93084; Q16115;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE SARCOPLASMIC/ENDOPLASMIC RETICULIUM CALCIUM ATPASE ISOFORM 3
 DE (EC 3.6.1.38) (CALCIUM PUMP 3) (SERCA3) (SR CA(2+)-ATPASE 3).
 GN ATP2A3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=LEUKEMIA;
 RX MEDLINE; 96404924.
 RA Dode L., Wuytack F., Kools P.F.J., Baba-Aissa F., Raeymaekers L.,
 RA Brike F., van de Ven W.J.M., Casteels R.;
 RA "cDNA cloning, expression and chromosomal localization of the human
 RA sarco/endoplasmic reticulum Ca(2+)-ATPase 3 gene.";
 RL Biochem. J. 318:689-699(1996).
 [2]
 RN ERRATUM.
 RA Dode L., Wuytack F., Kools P.F.J., Baba-Aissa F., Raeymaekers L.,
 RA Brike F., van de Ven W.J.M., Casteels R.;
 RL Biochem. J. 319:1008-1008(1996).
 [3]
 RN SEQUENCE OF 454-509 FROM N.A.
 RX MEDLINE; 94117459.
 RA Wuytack F., Papp B., Verboomen H., Raeymaekers L., Dode L.,
 RA Bohe R., Enouf J., Bokkila S., Authi K.S., Casteels R.;
 RA "A sarco/endoplasmic reticulum Ca(2+)-ATPase 3-type Ca2+ pump is
 RT expressed in platelets, in lymphoid cells, and in mast cells.";
 RL J. Biol. Chem. 269:1410-1416(1994).
 CC -1- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS
 CC OF ATP COUPLED WITH THE TRANSPORT OF THE CALCIUM. TRANSPORTS
 CC CALCIUM IONS FROM THE CYTOSOL INTO THE SARCOPLASMIC/ENDOPLASMIC
 CC RETICULUM. HAS A CENTRAL ROLE IN INTRACELLULAR CALCIUM SIGNALING.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O -> ADP + ORTHOPHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
 CC (E1-E2 ATPASES).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z69881; CAA93737.1; -;
 DR EMBL; Z69880; CAA93736.1; -;
 DR EMBL; S68239; AAB29700.1; -;
 DR MIM; 601929; -;
 DR INTERPRO; IPR001757; -;
 DR PFAM; PF00122; E1-E2_ATPase; 1.
 DR PRINTS; PR00119; CATATPASE.

DR PROSITE; PS00154; ATPASE_E1_E2; 1.
 KW Hydrolase; Calcium transport; Transmembrane; Phosphorylation;
 KW Magnesium; ATP-binding.
 FT DOMAIN 1 59
 FT TRANSMEM 60 78
 FT DOMAIN 79 86
 FT TRANSMEM 87 107
 FT DOMAIN 108 259
 FT TRANSMEM 298 315
 FT DOMAIN 316 772
 FT TRANSMEM 773 795
 FT DOMAIN 796 836
 FT TRANSMEM 837 859
 FT DOMAIN 860 894
 FT TRANSMEM 895 913
 FT DOMAIN 914 930
 FT TRANSMEM 931 950
 FT DOMAIN 951 999
 FT MOD_RES 351 351
 FT BINDING 515 515
 FT SEQUENCE 999 AA; 109237 MW; 162F333FF41BE071 CRC64;

Query Match 64.4%; Score 38; DB 1; Length 999;
 Best Local Similarity 70.0%; Pred. No. 31;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KLPRSPADSL 11
 :|:|:|:|:|:
 Db 819 KLPRSPREAL 828

RESULT 3

ID VP19_HCMVA STANDARD; PRT; 290 AA.
 AC P16783;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE PROBABLE CAPSID ASSEMBLY AND DNA MATURATION PROTEIN (CAPSID PROTEIN
 DE VP19C).
 GN UL46.
 OS Human cytomegalovirus (strain AD169).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE; 90269039.
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
 RA Horsnell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,
 RA Reddell E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
 RA "Analysis of the protein-coding content of the sequence of human
 RT cytomegalovirus strain AD169";
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
 CC -1- FUNCTION: COMPONENT OF THE BASAL LAYER IN WHICH THE CAPSIDS ARE
 CC EMBEDDED. BINDS DNA (BY SIMILARITY).
 CC -1- SIMILARITY: TO OTHER HERPESVIRUSES CAPSID PROTEIN VP19C.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X17403; CAA35405.1; -;
 DR PIR; S09809; S09809.
 KW Capsid assembly; Coat protein; DNA-binding.
 SQ SEQUENCE 290 AA; 33027 MW; ADB19CE379503D17 CRC64;

Query Match

61.0%; Score 36; DB 1; Length 290;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 09:03:55 ; Search time 58.45 seconds
(without alignments)
6.559 Million cell updates/sec

Title: US-09-236-468A-2_COPY_476_487
Perfect score: 59
Sequence: 1 AKLPRSPADSLT 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 87993

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	38	64.4	712	1	CM4C_HUMAN
2	38	64.4	999	1	ATC3_HUMAN
3	36	61.0	290	1	VP19_HCMVA
4	36	61.0	364	1	Y174_HUMAN
5	36	61.0	462	1	COGM_MOUSE
6	35	58.3	153	1	VPG_BYDVN
7	35	59.3	202	1	Y1BF_ECOLI
8	35	59.3	329	1	YCLC_PSEPU
9	35	59.3	331	1	CG12_HUMAN
10	35	59.3	383	1	APL_HUMAN
11	35	59.3	654	1	VPJ1_CAEEL
12	35	59.3	999	1	ATC3_RAT
13	34	57.6	481	1	Y209_ARCFU
14	34	57.6	491	1	Y039_HUMAN
15	34	57.6	501	1	YM05_YEAST
16	34	57.6	920	1	CO4_BOVIN
17	33	55.9	130	1	CYC3_DESYM
18	33	55.9	175	1	VPG_BMYVG
19	33	55.9	176	1	HMG4_SOYBN
20	33	55.9	240	1	SUR1_FUGRU
21	33	55.9	263	1	YG3S_YEAST
22	33	55.9	386	1	LMPL_EBV
23	33	55.9	404	1	LMPL_EBVC
24	33	55.9	420	1	MP13_MESAU
25	33	55.9	443	1	NU4M_CHLRE
26	33	55.9	571	1	TACY_STRPY
27	33	55.9	574	1	TACY_STRCB
28	33	55.9	574	1	TACY_STRQB
29	33	55.9	743	1	STB5_YEAST
30	33	55.9	836	1	NOT3_YEAST
31	33	55.9	855	1	CLOC_MOUSE
32	33	55.9	993	1	TSH_DROME
33	33	55.9	1045	1	SPS_BETVU

RESULT 1

ID	CM4C_HUMAN	STANDARD;	PRT;	712 AA.
AC	Q08493;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	CAMP-DEPENDENT 3',5'-CYCLIC PHOSPHODIESTERASE 4C (EC 3.1.4.17)			
DE	(DPDE1) (PDE21).			
GN	PDE4C.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A. (PDE4C1).			
RC	TISSUE-SUBSTANTIA NIGRA;			
RX	MEDLINE; 95145731.			
RA	Engels P., Sullivan M., Mueller T., Luebbert H.;			
RT	"Molecular cloning and functional expression in yeast of a human			
RT	CAMP-specific phosphodiesterase subtype (PDE IV-C).";			
RL	FEBS Lett. 358:305-310(1995).			
RN	[2]			
RP	SEQUENCE OF 462-712 FROM N.A.			
RX	MEDLINE; 94019330.			
RA	Bolger G., Michaeli T., Martins T., St John T., Steiner B.,			
RA	Rodgers L., Riggs M., Wigler M., Ferguson K.;			
RT	"A family of human phosphodiesterases homologous to the dunce			
RT	learning and memory gene product of Drosophila melanogaster are			
RT	potential targets for antidepressant drugs.";			
RL	Mol. Cell. Biol. 13:6558-6571(1993).			
CC	-!- CATALYTIC ACTIVITY: ADENOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O -			
CC	ADENOSINE 5'-PHOSPHATE.			
CC	-!- ENZYME REGULATION: INHIBITED BY ROLIPRAM.			
CC	-!- PATHWAY: CYCLIC NUCLEOTIDE METABOLISM.			
CC	-!- ALTERNATIVE PRODUCTS: VARIOUS ISOFORMS OF THE PROTEIN (PDE4C1 TO			
CC	PDE4C7) ARE PRODUCED BY ALTERNATIVE SPLICING. THE ISOFORM SHOWN			
CC	HERE IS PDE4C1.			
CC	-!- TISSUE SPECIFICITY: EXPRESSED IN VARIOUS TISSUES BUT NOT IN CELLS			
CC	OF THE IMMUNE SYSTEM.			
CC	-!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE			
CC	FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; Z46632; CAA86601.1; -			
DR	EMBL; L20968; AAA03591.1; -			
DR	MIM; 600128; -			
DR	INTERPRO; IPR002073; -			
DR	PFAM; PF00233; PDEase; 1.			
DR	PRINTS; PR00387; PD1ESTERASE1.			

075385 homo sapien
070405 mus musculus
P34609 caenorhabdi
P41233 mus musculus
Q61789 mus musculus
O28779 archaeoglob
P38133 streptomyce
O06235 mycobacteri
P37923 salmonella
O05837 mycobacteri
P03098 hamster pol
Q05424 neurospora

ALIGNMENTS

RESULT 1
ID CM4C_HUMAN STANDARD; PRT; 712 AA.

AC Q08493;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE CAMP-DEPENDENT 3',5'-CYCLIC PHOSPHODIESTERASE 4C (EC 3.1.4.17)

DE (DPDE1) (PDE21).

GN PDE4C.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]

RP SEQUENCE FROM N.A. (PDE4C1).

RC TISSUE-SUBSTANTIA NIGRA;

RX MEDLINE; 95145731.

RA Engels P., Sullivan M., Mueller T., Luebbert H.;

RT "Molecular cloning and functional expression in yeast of a human

CAMP-specific phosphodiesterase subtype (PDE IV-C).";

RL FEBS Lett. 358:305-310(1995).

RN [2]

RP SEQUENCE OF 462-712 FROM N.A.

RX MEDLINE; 94019330.

RA Bolger G., Michaeli T., Martins T., St John T., Steiner B.,

RA Rodgers L., Riggs M., Wigler M., Ferguson K.;

RT "A family of human phosphodiesterases homologous to the dunce

learning and memory gene product of Drosophila melanogaster are

potential targets for antidepressant drugs.";

RL Mol. Cell. Biol. 13:6558-6571(1993).

CC -!- CATALYTIC ACTIVITY: ADENOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O -

ADENOSINE 5'-PHOSPHATE.

CC -!- ENZYME REGULATION: INHIBITED BY ROLIPRAM.

CC -!- PATHWAY: CYCLIC NUCLEOTIDE METABOLISM.

CC -!- ALTERNATIVE PRODUCTS: VARIOUS ISOFORMS OF THE PROTEIN (PDE4C1 TO

PDE4C7) ARE PRODUCED BY ALTERNATIVE SPLICING. THE ISOFORM SHOWN

HERE IS PDE4C1.

CC -!- TISSUE SPECIFICITY: EXPRESSED IN VARIOUS TISSUES BUT NOT IN CELLS

OF THE IMMUNE SYSTEM.

CC -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE

FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; Z46632; CAA86601.1; -

DR EMBL; L20968; AAA03591.1; -

DR MIM; 600128; -

DR INTERPRO; IPR002073; -

DR PFAM; PF00233; PDEase; 1.

DR PRINTS; PR00387; PD1ESTERASE1.

THIS PAGE BLANK (USPTO)

OY 1 AKLPRSPADS 10
+ : | | | | | :
Db 116 AELERSPADA 125

Search completed: November 8, 2000, 08:53:36
Job time: 367 sec

C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T32228
R:Davidson, S.: Wohldmann, P.; Gillam, B.
submitted to the EMBL Data Library, September 1997
A:Description: The sequence of C. elegans cosmid T23B12.
A:Reference number: 221137
A:Accession: T32228
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1744 <DAV>
A:Cross-references: EMBL:AF022982; PIDN:AAB69941.1; GSPDB:GN00023; CESP:T23B12.9
A:Experimental source: strain Bristol N2; clone T23B12
C:Genetics:
A:Gene: CESP:T23B12.9
A:Map position: 5
A:Introns: 483/1; 646/1; 1241/1

Query Match 61.08; Score 36; DB 2; Length 1744;
Best Local Similarity 77.8%; Pred. No. 2.3e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PRSPADSLT 12
||:|||||
Db 1440 PRTPADSPT 1448

RESULT 12
T15962
hypothetical protein F07C7.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15962
R:Favetto, T.
submitted to the EMBL Data Library, January 1996
A:Description: The sequence of C. elegans cosmid F07C7.
A:Reference number: Z18437
A:Accession: T15962
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1879 <FAV>
A:Cross-references: EMBL:U46672; NID:g1166599; PID:g1166600; PIDN:AAA85753.1; CESP:F07C7
C:Genetics:
A:Gene: CESP:F07C7.1
A:Introns: 818/1; 1522/1; 1702/3; 1784/1

Query Match 61.08; Score 36; DB 2; Length 1879;
Best Local Similarity 77.8%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PRSPADSLT 12
||:|||||
Db 1606 PRTPADSPT 1614

RESULT 13
J01244
coat protein 2 - barley yellow dwarf virus
N:Alternate names: ORF5 protein
C:Species: barley yellow dwarf virus, BYDV
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 07-Apr-1994
C:Accession: J01244
R:Vincet, J.R.; Lister, R.M.; Larkins, B.A.
J. Gen. Virol. 72, 2347-2355, 1991
A:Title: Nucleotide sequence analysis and genomic organization of the NY-RPV isolate of
A:Reference number: J01240; MUID:92013950
A:Accession: J01244
A:Molecule type: genomic RNA
A:Residues: 1-152 <VIN>
A:Experimental source: strain NY-RPV
C:Comment: This virus causes yellowing disease in barley, oats, wheat, rice, maize and d

C:Superfamily: potato leaf roll virus genome-linked protein

Query Match 59.38; Score 35; DB 2; Length 152;
Best Local Similarity 70.08; Pred. No. 26;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LPRSPADSLT 12
||:|||||
Db 112 LPQQPAPSLT 121

RESULT 14
S47813
hypothetical protein b3592 - Escherichia coli
N:Alternate names: hypothetical protein f202
C:Species: Escherichia coli
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 24-Nov-1999
C:Accession: S47813; B65159
R:Plunkett, G.
submitted to the EMBL Data Library, March 1994
A:Reference number: S47666
A:Accession: S47813
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-202 <PLU>
A:Cross-references: EMBL:U00039; NID:q466582; PIDN:AAB18569.1; PID:q466730
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: B65159
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-202 <BLAT>
A:Cross-references: GB:AE000436; GB:U00096; NID:g2367246; PIDN:AAC76616.1; PID:g17900
A:Experimental source: strain K-12, substrain MG1655
C:Superfamily: auxin-induced protein

Query Match 59.38; Score 35; DB 2; Length 202;
Best Local Similarity 63.6%; Pred. No. 35;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AKLPRSPADSL 11
|||:|:|
Db 78 AMLPRDPLESL 88

RESULT 15
T35041
hypothetical protein SC4G2.06 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 04-Mar-2000
C:Accession: T35041
R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1998
A:Reference number: 221566
A:Accession: T35041
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-209 <SSE>
A:Cross-references: EMBL:AL031371; PIDN:CAA20543.1; GSPDB:GN00070; SCOEDB:SC4G2.06
C:Genetics:
A:Experimental source: strain A3(2)
C:Superfamily: Streptomyces coelicolor hypothetical protein SC4G2.06

Query Match 59.38; Score 35; DB 2; Length 209;
Best Local Similarity 70.08; Pred. No. 36;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

A:Residues: 1-1407 <PAR>
A:Cross-references: EMBL:AL023496; PIDN:CAA18919.1
C:Superfamily: enterobactin synthetase component F; acetate--CoA ligase homology; acyl carrier protein homology
C:Keywords: carrier protein; phosphopantetheine
F:532-1145/Domain: gramicidin S synthetase I repeat homology <GRS>
F:591-1049/Domain: acetate--CoA ligase homology <ACL>
F:1077-1145/Domain: acyl carrier protein homology <ACP>
F:1109/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 64.4%; Score 38; DB 2; Length 1407;
Best Local Similarity 77.8%; Pred. No. 79;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LPRSPADSL 11
||| ||| |||
DB 1306 LPREPADAL 1314

RESULT 7
T47705
hypothetical protein F1116.70 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Jun-2000
C:Accession: T47705
submitted to the Protein Sequence Database, March 2000
R:Benes, V.; Wurmbach, E.; Drzonek, H.; Ansoerge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.H.
A:Reference number: 224473
A:Accession: T47705
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-579 <BEN>
A:Cross-references: EMBL:AL161667
A:Experimental source: cultivar Columbia; BAC clone F1116
C:Genetics:
A:Map position: 3
A:Introns: 107/1; 146/1; 197/3; 222/3; 315/3; 437/3
A:Note: F1116.70
C:Superfamily: Arabidopsis thaliana hypothetical protein T7H20.60

Query Match 62.7%; Score 37; DB 2; Length 579;
Best Local Similarity 70.0%; Pred. No. 46;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LPRSPADSLT 12
||| ||| |||
DB 262 LPRVPSDGLT 271

RESULT 8
S09809
hypothetical protein UL46 - human cytomegalovirus (strain AD169)
C:Species: human cytomegalovirus, human herpesvirus 5
A:Note: host Homo sapiens (man)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Oct-1999
C:Accession: S09809
R:Chen, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.; M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A:Reference number: S09749; MUID:90269039
A:Accession: S09809
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-290 <CHE>
A:Cross-references: EMBL:X17403; NID:g59591; PIDN:CAA35405.1; PID:g1780824
A:Note: this sequence was submitted to the EMBL Data Library, December 1989

Query Match 61.0%; Score 36; DB 2; Length 290;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKLPRSPAD 9
||| ||| |||
DB 7 AKRPRDPAD 15

RESULT 9
T32184
hypothetical protein ZK994.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T32184
R:Davidson, S.; Wohldmann, P.
A:Description: The sequence of C. elegans cosmid ZK994.
A:Reference number: Z21133
A:Accession: T32184
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-339 <DAV>
A:Cross-references: EMBL:AF022977; PIDN:AA88613.1; GSPDB:GN00023; CESP:ZK994.5
A:Experimental source: strain Bristol N2; clone ZK994
C:Genetics:
A:Gene: CESP:ZK994.5
A:Map position: 5

Query Match 61.0%; Score 36; DB 2; Length 339;
Best Local Similarity 77.8%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PRSPADSLT 12
||| ||| |||
DB 35 PRTPADSPST 43

RESULT 10
A42401
macrophage elastase (EC 3.4.24.-) precursor - mouse
N:Alternate names: matrix metalloproteinase 12 (MMP12)
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: A42401
R:Shapiro, S.D.; Griffin, G.L.; Gilbert, D.J.; Jenkins, N.A.; Copeland, N.G.; Welgus, J. Biol. Chem. 267, 4664-4671, 1992
A:Title: Molecular cloning, chromosomal localization, and bacterial expression of a matrix metalloproteinase gene from mouse
A:Reference number: A42401; MUID:92163826
A:Accession: A42401
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-462 <SHA>
A:Cross-references: GB:M82831; NID:g199127; PIDN:AAA39526.1; PID:g199128
C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metallopro
C:Keywords: hydrolase; metalloproteinase; zinc; zymogen
F:53-256/Domain: matrix metalloproteinase homology <MMP>
F:269-462/Domain: matrix metalloproteinase homology <PXA>
F:85-211,215,221/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #stat
F:211,215,221/Binding site: zinc, catalytic (His) (active) #status predicted
F:212/Active site: Glu #status predicted

Query Match 61.0%; Score 36; DB 2; Length 462;
Best Local Similarity 63.6%; Pred. No. 56;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 KLPRSPADSLT 12
||| ||| |||
DB 302 KLPGSPATNIT 312

RESULT 11
T32228
hypothetical protein T23B12.9 - Caenorhabditis elegans

Db 4 AKLPRSPSE 12

RESULT 3

I61356
phosphodiesterase - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 09-Jun-2000
C:Accession: I61356
R:Bolger, G.; Michaeli, T.; Martins, T.; St. John, T.; Steiner, B.; Rodgers, L.; Riggs, M.L. Cell. Biol. 13, 6558-6571, 1993
A:Title: A family of human phosphodiesterases homologous to the dunce learning and memory gene
A:Reference number: A54442; MUID:94019330
A:Accession: I61356
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-251 <RES>
A:Cross-references: GB:I20968; NID:q347125; PIDN:AAA03591.1; PID:q347126
F:1-154/Domain: 3',5'-cyclic-nucleotide phosphodiesterase homology <CNP>

Query Match 64.4%; Score 38; DB 2; Length 251;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9

Db 175 SKIPRSPD 183

RESULT 4

S71626
3',5'-cyclic-nucleotide phosphodiesterase (EC 3.1.4.17), cAMP-specific - human
N:Alternate names: 3',5'-cyclic AMP phosphodiesterase
C:Species: Homo sapiens (man)
C:Date: 27-Nov-1997 #sequence_revision 12-Dec-1997 #text_change 09-Jun-2000
C:Accession: S71626
R:Engels, P.; Sullivan, M.; Mueller, T.; Luebbert, H.
FEBS Lett. 358, 305-310, 1995

A:Title: Molecular cloning and functional expression in yeast of a human cAMP-specific phosphodiesterase
A:Reference number: S71626; MUID:95145731

A:Accession: S71626
A:Molecule type: DNA
A:Residues: 1-712 <ENG>
A:Cross-references: EMBL:246632; NID:g727222; PID:g727223
A:Experimental source: substantia nigra

C:Genetics:
A:Gene: HSPDE4C1

C:Function:
A:Description: cAMP hydrolysis; converts cAMP to non-cyclic AMP

A:Pathway: cyclic nucleotide metabolism
A:Note: expressed in various tissues but not in cells of the immune system
C:Superfamily: 3',5'-cyclic-nucleotide phosphodiesterase homology
C:Keywords: phosphoric diester hydrolase
F:367-615/Domain: 3',5'-cyclic-nucleotide phosphodiesterase homology <CNP>

Query Match 64.4%; Score 38; DB 2; Length 712;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9

Db 636 SKIPRSPD 644

RESULT 5

S72267
Ca2+-transporting ATPase (EC 3.6.1.38) isoform SERCA3, sarcoplasmic/endoplasmic reticulum
C:Species: Homo sapiens (man)
C:Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 20-Jun-2000
C:Accession: S72267; S72268
R:Dode, L.; Wuytack, F.; Kools, P.F.J.; Baba-Aissa, F.; Raeymaekers, L.; Brik, F.; van d

Biochem. J. 318, 689-699, 1996
A:Title: cDNA cloning, expression and chromosomal localization of the human sarco/endoplasmic reticulum Ca²⁺-ATPase isoform SERCA3
A:Reference number: S72267; MUID:96404924
A:Accession: S72267
A:Molecule type: mRNA
A:Residues: 1-999 <DOD>
A:Cross-references: EMBL:Z69881; NID:gl524091; PIDN:CAA93737.1; PID:gl524092
A:Experimental source: sarco/endoplasmic reticulum
A:Accession: S72268
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 916-999 <DOW>
A:Cross-references: EMBL:Z69880; NID:gl524089; PIDN:CAA93736.1; PID:gl524090
A:Experimental source: sarco/endoplasmic reticulum
A:Note: list of introns is incomplete
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996

C:Genetics:

A:Gene: SERCA3
A:Map position: 17p13.3
A:Introns: 954/3; 994/1

C:Function:

A:Description: catalyzes hydrolysis of one molecule ATP coupled to translocation of two Ca²⁺ ions across the sarcoplasmic or endoplasmic reticulum; Ca²⁺ pump
A:Note: membrane-bound enzyme; magnesium-dependent; P-type ATPase
C:Superfamily: Na⁺/K⁺-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: alternative splicing; ATP; calcium binding; calcium transport; endoplasmic reticulum; hydrolysis
F:40-57/Domain: calcium binding #status predicted <CA1>
F:60-78/Domain: transmembrane #status predicted <TM01>
F:87-107/Domain: transmembrane #status predicted <TM02>
F:108-257/Domain: intracellular #status predicted <INT1>
F:111-131/Domain: calcium binding #status predicted <CA2>
F:132-238/Domain: transduction #status predicted <TSD>
F:238-256/Domain: calcium binding #status predicted <CA3>
F:263-279/Domain: transmembrane #status predicted <TM03>
F:298-316/Domain: transmembrane #status predicted <TM04>
F:308-760/Domain: intracellular #status predicted <INT2>
F:310-329/Domain: calcium binding #status predicted <CA4>
F:330-505/Domain: catalytic #status predicted <PHY>
F:506-681/Domain: ATP binding #status predicted <ATP>
F:595-768/Domain: ATPase nucleotide-binding domain homology <ATN>
F:681-738/Domain: hinge #status predicted <HNG>
F:739-760/Domain: calcium binding #status predicted <CA5>
F:763-784/Domain: transmembrane #status predicted <TM05>
F:788-809/Domain: transmembrane #status predicted <TM06>
F:837-857/Domain: transmembrane #status predicted <TM07>
F:894-913/Domain: transmembrane #status predicted <TM08>
F:931-950/Domain: transmembrane #status predicted <TM09>
F:959-980/Domain: transmembrane #status predicted <TM10>
F:351/Active site: Asp (aspartylphosphate intermediate) #status predicted
F:515/Binding site: ATP (Lys) #status predicted

Query Match 64.4%; Score 38; DB 2; Length 999;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KLPRSPADSL 11

Db 819 KLPRSPREAL 828

RESULT 6

T28702
probable polyketide synthetase [similarity] - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T28702
R:Parkhill, J.; Bentley, S.D.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, January 1999
A:Reference number: Z20512
A:Accession: T28702
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:53:34 ; Search time 99.87 Seconds
(without alignments)
7.625 Million cell updates/sec

Title: US-09-236-468a-2_COPY_476_487

Perfect score: 59

Sequence: 1 AKLPRSPADSLT 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 182106

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_65:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	66.1	182	3 JC7196	phytoecyanin-relate
2	39	66.1	1237	2 T46609	calcium-activated
3	38	64.4	251	2 I61356	phosphodiesterase
4	38	64.4	712	2 S71626	3',5'-cyclic-nucle
5	38	64.4	999	2 S72267	Ca2+-transporting
6	38	64.4	1407	2 T28702	probable polyketid
7	37	62.7	579	2 T47705	hypothetical prote
8	36	61.0	290	2 S09809	hypothetical prote
9	36	61.0	339	2 T32184	hypothetical prote
10	36	61.0	462	2 A42401	macrophage elastat
11	36	61.0	1744	2 T32228	hypothetical prote
12	36	61.0	1879	2 T15962	hypothetical prote
13	35	59.3	152	2 J01244	coat protein 2 - b
14	35	59.3	202	2 S47813	hypothetical prote
15	35	59.3	209	2 T35041	hypothetical prote
16	35	59.3	281	2 T44620	conserved hypot het
17	35	59.3	282	2 T44668	hypothetical prote
18	35	59.3	284	2 T75418	citrate lyase, bet
19	35	59.3	349	2 T43457	hypothetical prote
20	35	59.3	356	2 T40041	hypothetical prote
21	35	59.3	654	2 T19785	hypothetical prote
22	35	59.3	885	2 S67660	hypothetical prote
23	35	59.3	999	2 A34307	Ca2+-transporting
24	34	57.6	287	2 C72099	conserved hypot het
25	34	57.6	316	2 B70571	hypothetical prote
26	34	57.6	336	2 H72618	hypothetical prote
27	34	57.6	381	2 I76776	Mannosyltransferas
28	34	57.6	479	2 T23508	hypothetical prote
29	34	57.6	481	2 A69276	conserved hypot het

ALIGNMENTS

RESULT 1

JC7196

phytoecyanin-related protein Pn14 - common morning-glory

C:Species: Ipomoea purpurea (common morning-glory)

C>Date: 15-Mar-2000 #sequence_revision 15-Mar-2000 #text_change 15-Mar-2000

C:Accession: JC7196

R:Yoshizaki, M.; Furumoto, T.; Hata, S.; Shinozaki, M.; Izui, K.

Biochem. Biophys. Res. Commun. 268, 466-470, 2000

A:Title: Characterization of a novel gene encoding a phytoecyanin-related protein in m

A:Reference number: JC7196; MUID:20145457

A:Accession: JC7196

A>Status: preliminary

A:Residues: 1-182 <YOS>

A:Cross-references: DDBJ:AB035146

Query Match 66.1%; Score 39; DB 3; Length 182;

Best Local Similarity 70.0%; Pred. No. 5.7;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 2 KLPRSPADSL 11

Db 39 KIPSPSDSL 48

RESULT 2

T46609

calcium-activated potassium channel protein - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 31-Mar-2000

C:Accession: T46609

R:Joiner, W.J.; Tang, M.D.; Wang, L.Y.; Dworetzky, S.I.; Boissard, C.G.; Gan, L.; Gri

Nature Neurosci. 1, 462-469, 1998

A:Title: Formation of intermediate-conductance calcium-activated potassium channels b

A:Reference number: 223103; MUID:99212276

A:Accession: T46609

A>Status: translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-1237 <JOI>

A:Cross-references: EMBL:AF089730; NID:g3978471; PID:g3978472; PIDN:AAC83350.1

A>Note: this publication cited in GenBank entry AF089730 (not Medline), release 114,

C:Genetics:

A:Gene: Slack

A>Note: abundantly expressed in the nervous system

Query Match 66.1%; Score 39; DB 2; Length 1237;

Best Local Similarity 77.8%; Pred. No. 4.5;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AKLPRSPAD 9

|||||:

THIS PAGE BLANK (USPTO)

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/771.602D
; FILING DATE: 20-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/010,869
; FILING DATE: 31-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 8-96
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 94 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..94
; OTHER INFORMATION: /note= "Xaa residues are not
; OTHER INFORMATION: specified in this sequence."
US-08-771-602D-49

Query Match 55.9%; Score 33; DB 2; Length 94;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 5 RSPADSLT 12
|:|:|:|:
Db 84 RNPADALT 91

Search completed: November 8, 2000, 08:49:23
Job time: 116 sec

; SEQUENCE CHARACTERISTICS:
; LENGTH: 506 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-286-856C-3

Query Match 64.4%; Score 38; DB 1; Length 506;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
Db 430 SKIPRSPSD 438

RESULT 13
US-08-472-831-3
; Sequence 3, Application US/08472831
; Patent No. 5686286
; GENERAL INFORMATION:
; APPLICANT: Fisher, Douglas A.
; TITLE OF INVENTION: hrDE IV-C: A NOVEL HUMAN
; TITLE OF INVENTION: PHOSPHODIESTERASE IV ISOZYME
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pfizer Inc.
; STREET: 235 East 42nd Street, 20th Floor
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10017-5755
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,831
; FILING DATE: June 7, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sheyka, Robert F.
; REGISTRATION NUMBER: 31,304
; REFERENCE/DOCKET NUMBER: PC8552B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)573-5425
; TELEFAX: (212)573-1939
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 506 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-472-831-3

Query Match 64.4%; Score 38; DB 1; Length 506;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
Db 430 SKIPRSPSD 438

RESULT 14
US-08-577-492-32
; Sequence 32, Application US/08577492
; Patent No. 5851784

; GENERAL INFORMATION:
; APPLICANT: Owens, Raymond John
; APPLICANT: Perry, Martin John
; APPLICANT: Lumb, Simon Mark
; TITLE OF INVENTION: HUMAN PHOSPHODIESTERASE TYPE IVC, AND
; TITLE OF INVENTION: ITS PRODUCTION AND USE
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5851784ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/577,492
; FILING DATE: 22-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9426227.6
; FILING DATE: 23-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9512996.1
; FILING DATE: 26-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cherry, David A.
; REGISTRATION NUMBER: 35,099
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 606 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
US-08-577-492-32

Query Match 64.4%; Score 38; DB 2; Length 606;
Best Local Similarity 66.7%; Pred. No. 34;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
Db 530 SKIPRSPSD 538

RESULT 15
US-08-771-602D-49
; Sequence 49, Application US/08771602D
; Patent No. 5976795
; GENERAL INFORMATION:
; APPLICANT: Voytas, Daniel F.
; APPLICANT: Zou, Sige
; TITLE OF INVENTION: Retrotransposon and Methods
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

QY 1 AKLPRSPAD 9
:|:|||||
Db 405 SKIPRSPD 413

RESULT 10
US-08-472-831-2
; Sequence 2, Application US/08472831
; Patent No. 5686286
; GENERAL INFORMATION:
; APPLICANT: Fisher, Douglas A.
; TITLE OF INVENTION: hPDE IV-C: A NOVEL HUMAN
; TITLE OF INVENTION: PHOSPHODIESTERASE IV ISOZYME
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pfizer Inc.
; STREET: 235 East 42nd Street, 20th Floor
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10017-5755
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,831
; FILING DATE: June 7, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sheyka, Robert F.
; REGISTRATION NUMBER: 31,304
; REFERENCE/DOCKET NUMBER: PC8552B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)573-5425
; TELEFAX: (212)573-1939
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-472-831-2

Query Match 64.4%; Score 38; DB 1; Length 481;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLPRSPAD 9
:|:|||||
Db 405 SKIPRSPD 413

RESULT 11
US-08-577-492-40
; Sequence 40, Application US/08577492
; Patent No. 5851784
; GENERAL INFORMATION:
; APPLICANT: Owens, Raymond John
; APPLICANT: Perry, Martin John
; APPLICANT: Lumb, Simon Mark
; TITLE OF INVENTION: HUMAN PHOSPHODIESTERASE TYPE IVC, AND
; TITLE OF INVENTION: ITS PRODUCTION AND USE
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5851784rls
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA

; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/577,492
; FILING DATE: 22-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9426227.6
; FILING DATE: 23-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9512996.1
; FILING DATE: 26-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cherry, David A.
; REGISTRATION NUMBER: 35,099
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 501 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
US-08-577-492-40

Query Match 64.4%; Score 38; DB 2; Length 501;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLPRSPAD 9
:|:|||||
Db 425 SKIPRSPD 433

RESULT 12
US-08-286-856C-3
; Sequence 3, Application US/08286856C
; Patent No. 5672509
; GENERAL INFORMATION:
; APPLICANT: FISHER, DOUGLAS A
; TITLE OF INVENTION: hPDE IV-C: A NOVEL HUMAN
; TITLE OF INVENTION: PHOSPHODIESTERASE IV
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PFIZER INC
; STREET: 235 EAST 42ND STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10017-5755
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,856C
; FILING DATE: 05-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SHEYKA, ROBERT F
; REGISTRATION NUMBER: 31304
; REFERENCE/DOCKET NUMBER: PC8552A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-573-1189
; TELEFAX: 212-573-1939
; INFORMATION FOR SEQ ID NO: 3:

MACCLES, V., CONSER

6

Query Match 64.4%; Score 38; DB 1; Length 481;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 6; Conservative 3; Mismatches 0; Indels

REFERENCE/DOCKET NUMBER: 27805/30197
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 222 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US91-02714-43

Query Match 64.4%; Score 38; DB 4; Length 222;
Best Local Similarity 66.7%; Pred. NO. 11;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
Db 146 SKIPRSPD 154

RESULT 5
PCT-US91-02714-54
Sequence 54, Application PC/TUS9102714
GENERAL INFORMATION:
APPLICANT: Wigler, Michael H.
APPLICANT: Colicelli, John J.
TITLE OF INVENTION: Cloning by Complementation and Related
TITLE OF INVENTION: Processes
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Bicknell
STREET: Two First National Plaza, 20 South Clark
STREET: Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02714
FILING DATE: 19910419
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25447
REFERENCE/DOCKET NUMBER: 27805/30197
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 222 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US91-02714-54

Query Match 64.4%; Score 38; DB 4; Length 222;
Best Local Similarity 66.7%; Pred. NO. 11;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
Db 146 SKIPRSPD 154
RESULT 6
US-08-474-379C-46
Sequence 46, Application US/08474379C
Patent No. 5977305
GENERAL INFORMATION:
APPLICANT: Wigler, Michael H.
APPLICANT: Colicelli, John J.
TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED
TITLE OF INVENTION: PROCESSES
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,379C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/206,188
FILING DATE: 01-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/688,352
FILING DATE: 19-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 27866/32771
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-379C-46

Query Match 64.4%; Score 38; DB 2; Length 253;
Best Local Similarity 66.7%; Pred. NO. 13;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
Db 177 SKIPRSPD 185

RESULT 7
US-09-146-249A-46
Sequence 46, Application US/09146249A
Patent No. 6069240
GENERAL INFORMATION:
APPLICANT: Wigler, Michael H.
APPLICANT: Colicelli, John J.

Db 476 AKLPRSPADSLT 487

RESULT 2

PCT-US95-07085-2

Sequence 2, Application PC/TUS9507085

GENERAL INFORMATION:

APPLICANT: Soppet, Daniel R

APPLICANT: Yi, Li

APPLICANT: Rosen, Craig A

APPLICANT: Ruben, Steven

TITLE OF INVENTION: G-Protein Parathyroid Hormone Receptor

TITLE OF INVENTION: HLTC74

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,

ADDRESSEE: Stewart & Olstein

STREET: 6 Becker Farm Road

CITY: Roseland

STATE: NJ

COUNTRY: USA

ZIP: 07068-1739

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/07085

FILING DATE: 05-JUN-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.

REGISTRATION NUMBER: 33,073

REFERENCE/DOCKET NUMBER: 325800-393

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 541 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US95-07085-2

Query Match 100.0%; Score 59; DB 4; Length 541;

Best Local Similarity 100.0%; Pred. No. 0.0055;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLPRSPADSLT 12

Db 476 AKLPRSPADSLT 487

RESULT 3

US-07-688-352C-46

Sequence 46, Application US/07688352C

Patent No. 5527896

GENERAL INFORMATION:

APPLICANT: Wigler, Michael H.

APPLICANT: Colicelli, John J.

TITLE OF INVENTION: Cloning by Complementation and Related

TITLE OF INVENTION: Processes

NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

ADDRESSEE: Bicknell

STREET: Two First National Plaza, 20 South Clark

STREET: Street

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/02714

FILING DATE: 19910419

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/511,715

FILING DATE: 20-APR-1990

ATTORNEY/AGENT INFORMATION:

NAME: Borun, Michael F.

REGISTRATION NUMBER: 25447

Query Match 64.4%; Score 38; DB 1; Length 222;

Best Local Similarity 66.7%; Pred. No. 11;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLPRSPAD 9

Db 146 SKIPRSPSD 154

RESULT 4

PCT-US91-02714-43

Sequence 43, Application PC/TUS9102714

GENERAL INFORMATION:

APPLICANT: Wigler, Michael H.

APPLICANT: Colicelli, John J.

TITLE OF INVENTION: Cloning by Complementation and Related

TITLE OF INVENTION: Processes

NUMBER OF SEQUENCES: 55

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

ADDRESSEE: Bicknell

STREET: Two First National Plaza, 20 South Clark

STREET: Street

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/02714

FILING DATE: 19910419

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/511,715

FILING DATE: 20-APR-1990

ATTORNEY/AGENT INFORMATION:

NAME: Borun, Michael F.

REGISTRATION NUMBER: 25447

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:49:22 ; Search time 97.15 Seconds
(without alignments)
2.070 Million cell updates/sec

Title: US-09-236-468A-2_COPY_476_487

Perfect score: 59

Sequence: 1 AKLPRSPADSLT 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 164575

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_AA:*
- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
 - 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
 - 3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
 - 4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
 - 5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	59	100.0	541	3	US-08-468-011A-2
2	59	100.0	541	4	PCT-US93-07085-2
3	38	64.4	222	1	US-07-688-352C-46
4	38	64.4	222	4	PCT-US91-02714-43
5	38	64.4	222	4	PCT-US91-02714-54
6	38	64.4	253	2	US-08-474-379C-46
7	38	64.4	253	3	US-09-146-249A-46
8	38	64.4	253	3	US-08-206-188B-46
9	38	64.4	481	1	US-08-286-856C-2
10	38	64.4	481	2	US-08-472-831-2
11	38	64.4	501	2	US-08-577-492-40
12	38	64.4	506	1	US-08-286-856C-3
13	38	64.4	506	1	US-08-472-831-3
14	38	64.4	606	2	US-08-577-492-32
15	33	55.9	94	2	US-08-771-602D-49
16	33	55.9	404	2	US-08-504-459-8
17	33	55.9	480	1	US-07-752-429E-2
18	33	55.9	480	1	US-07-752-428C-2
19	33	55.9	480	1	US-07-752-428C-4
20	33	55.9	855	2	US-08-816-693A-2
21	33	55.9	855	3	US-08-885-291-2
22	33	55.9	1045	2	US-08-553-436A-6
23	32	54.2	230	1	US-08-597-236-4
24	32	54.2	230	1	US-08-746-682A-4
25	32	54.2	265	2	US-08-898-779-2
26	32	54.2	265	2	US-08-898-779-4
27	32	54.2	359	1	US-07-914-281-14
28	32	54.2	359	1	US-08-393-246-14

29	32	54.2	359	1	US-08-525-058A-14	Sequence 14, Appl
30	32	54.2	359	2	US-08-696-731-14	Sequence 14, Appl
31	32	54.2	432	1	US-08-167-919A-11	Sequence 11, Appl
32	32	54.2	432	3	US-08-715-106-11	Sequence 11, Appl
33	32	54.2	461	1	US-08-385-229-2	Sequence 2, Appl
34	32	54.2	461	2	US-08-650-000-2	Sequence 2, Appl
35	32	54.2	461	5	5395760-2	Patent No. 5395760
36	32	54.2	507	4	PCT-US93-08386-8	Sequence 8, Appl
37	32	54.2	599	2	US-08-426-125-1	Sequence 1, Appl
38	32	54.2	599	2	US-08-426-125-3	Sequence 3, Appl
39	32	54.2	599	2	US-08-455-355-1	Sequence 1, Appl
40	32	54.2	599	2	US-08-455-355-3	Sequence 3, Appl
41	32	54.2	638	2	US-08-426-125-4	Sequence 4, Appl
42	32	54.2	638	2	US-08-455-355-4	Sequence 4, Appl
43	31	52.5	119	1	US-08-417-460-8	Sequence 8, Appl
44	31	52.5	183	3	US-09-122-443-11	Sequence 11, Appl
45	31	52.5	192	1	US-07-749-446-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-468-011A-2
; Sequence 2, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-011A-2

Query Match 100.0%; Score 59; DB 3; Length 541;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AKLPRSPADSLT 12
|||||

Oy 2 KLPRSPADSLT 12
: | | | | | | |
Db 765 rappepadslt 775

Search completed: November 8, 2000, 08:51:50
Job time: 262 sec

CC Huntingtin interacting protein HIP1 (W18030) is a brain-specific
CC protein that interacts differently with the gene product of a
CC normal (16 CAG repeat) and an expanded (over 44 CAG repeat)
CC Huntingtin's disease (HD) gene, and which facilitates incorporation
CC of huntingtin into brain cell membranes. Its amino acid sequence
CC was deduced from a human frontal cortex cDNA clone (f67187). The
CC effects of HD in a patient expressing huntingtin protein with an
CC expanded polyglutamine tract can be ameliorated by increasing the
CC amount of expressed HIP1 in the brain using gene therapy
CC approaches. Modified forms of HIP1 which bind more effectively to
CC expanded HD protein can be used to convert the expanded protein
CC into a functional molecule.
XX
SQ Sequence 914 AA;

Query Match 61.0%; Score 36; DB 18; Length 914;
Best Local Similarity 63.6%; Pred. No. 2e+02; 3; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KLPRSPADSLT 12
: | |||||
Db 589 rappepadslt 599

RESULT 14
Y59269
ID Y59269 standard; Protein; 914 AA.
XX
AC Y59269;
XX
DT 17-APR-2000 (first entry)
XX
DE Human huntingtin-interacting protein (HIP1).
XX
KW Huntingtin Interacting Protein; HIP; death effector domain; DED; human;
KW apoptosis; HIP-apoptosis modulating protein; cell death; gene therapy;
KW Huntington's disease; neurotropic; anticonvulsant; cytostatic;
KW chromosome 7q11.23.
XX
OS Homo sapiens.
XX
PN W09960986-A2.
XX
PD 02-DEC-1999.
XX
PF 27-MAY-1999; 99WO-US11743.
XX
PR 27-MAY-1998; 98US-0085199.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (MERI) MERCK FROSST CANADA INC.
XX
PI Kalchman M, Hayden MR, Hackam A, Chopra V, Nicholson DW;
PI Vallaincourt JP, Rasper DM;
XX
DR WPI: 2000-097055/08.
XX
PR 27-MAY-1998; 98US-0085199.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (MERI) MERCK FROSST CANADA INC.
XX
PI Kalchman M, Hayden MR, Hackam A, Chopra V, Nicholson DW;
PI Vallaincourt JP, Rasper DM;
XX
DR WPI: 2000-097055/08.
XX
PR N-PSDB; 258746.
XX
PT Novel proteins useful for treating Huntington's disease by gene therapy
XX techniques, and cancers -
XX
PS Claim 9; Page 53-57; 91pp; English.
XX
XX The invention relates to Huntingtin Interacting Protein (HIP), that
XX includes a death effector domain (DED), suggesting apoptotic function.
XX Proteins with DED (referred as HIP-apoptosis modulating proteins) are
XX useful for inducing apoptotic death in cells. The HIP is a logical target
XX for therapy in Huntington's disease since it has a differential apoptotic
XX activity, modulated by interaction with Huntingtin having normal and
XX expanded repeats. HIP is also used as a therapeutic agent to introduce
XX apoptosis in cancer cells. Increasing expression of normal (non-expanded)
XX Huntingtin or the HIP-apoptotic modulator-binding portion, a modified

CC HIP-apoptotic modulator in which the DED has been deleted, is helpful for
CC treating Huntington's disease. The present sequence represents the human
CC HIP1.
XX
SQ Sequence 914 AA;

Query Match 61.0%; Score 36; DB 21; Length 914;
Best Local Similarity 63.6%; Pred. No. 2e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KLPRSPADSLT 12
: | |||||
Db 589 rappepadslt 599

RESULT 15
Y59270
ID Y59270 standard; Protein; 1090 AA.
XX
AC Y59270;
XX
DT 17-APR-2000 (first entry)
XX
DE Human huntingtin-interacting protein (HIP1).
XX
KW Huntingtin Interacting Protein; HIP; death effector domain; DED; human;
KW apoptosis; HIP-apoptosis modulating protein; cell death; gene therapy;
KW Huntington's disease; neurotropic; anticonvulsant; cytostatic;
KW chromosome 7q11.23.
XX
OS Homo sapiens.
XX
PN W09960986-A2.
XX
PD 02-DEC-1999.
XX
PF 27-MAY-1999; 99WO-US11743.
XX
PR 27-MAY-1998; 98US-0085199.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (MERI) MERCK FROSST CANADA INC.
XX
PI Kalchman M, Hayden MR, Hackam A, Chopra V, Nicholson DW;
PI Vallaincourt JP, Rasper DM;
XX
DR WPI: 2000-097055/08.
XX
PT Novel proteins useful for treating Huntington's disease by gene therapy
XX techniques, and cancers -
XX
PS Claim 1; Page 57-62; 91pp; English.
XX
XX The invention relates to Huntingtin Interacting Protein (HIP), that
XX includes a death effector domain (DED), suggesting apoptotic function.
XX Proteins with DED (referred as HIP-apoptosis modulating proteins) are
XX useful for inducing apoptotic death in cells. The HIP is a logical target
XX for therapy in Huntington's disease since it has a differential apoptotic
XX activity, modulated by interaction with Huntingtin having normal and
XX expanded repeats. HIP is also used as a therapeutic agent to introduce
XX apoptosis in cancer cells. Increasing expression of normal (non-expanded)
XX Huntingtin or the HIP-apoptotic modulator-binding portion, a modified
XX HIP-apoptotic modulator in which the DED has been deleted, is helpful for
XX treating Huntington's disease. The present sequence represents the human
XX HIP1.
XX
SQ Sequence 1090 AA;

Query Match 61.0%; Score 36; DB 21; Length 1090;
Best Local Similarity 63.6%; Pred. No. 2.4e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

DE Human phosphodiesterase type IV C.
XX Phosphodiesterase; screening; identification; inhibitor;
KW inhibition; PDE; treatment; prophylaxis; inflammatory disease;
KW inflamed lung; asthma.
XX Homo sapiens.
OS
XX WO9620281-A1.
PN
XX
PD 04-JUL-1996.
XX
XX 21-DEC-1995; 95WO-GB03006.
PF
XX 26-JUN-1995; 95GB-0012996.
PR
XX 23-DEC-1994; 94GB-0026227.
PR
XX (CLLT) CELLTech THERAPEUTICS LTD.
PA
XX Lumb SM, Owens RJ, Perry MJ;
PI
XX WPI; 1996-321854/32.
DR
XX Human phosphodiesterase type IVC and selective inhibitors - used in
PT the treatment of inflammatory disease, esp. asthma.
PT
XX
PS Disclosure: Figure 2; 50pp; English.
XX
XX Recombinant phosphodiesterase (PDE) type IVC may be used to screen
CC for inhibitors of PDE IVC. The inhibitors may be used in
CC pharmaceutical for the treatment and prophylaxis of inflammatory
CC diseases, especially inflamed lung associated with asthma. Multiple
CC isoforms of PDE exist opening the possibility for individual
CC inhibitors of each isoform.
XX
XX Sequence 606 AA;
SQ

Query Match 64.4%; Score 38; DB 17; Length 606;
Best Local Similarity 66.7%; Pred. No. 56;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Oy 1 AKLPRSPAD 9
:|:||||:|
Db 530 skiprpsd 538

RESULT 12
Y59272
ID Y59272 standard; Protein; 756 AA.
XX
XX Y59272;
AC
XX
DT 17-APR-2000 (first entry)
XX
XX Mouse huntingtin-interacting protein (mH1P1).
DE
XX Huntingtin Interacting Protein; HIP; death effector domain; DEP; human;
KW apoptosis; HIP-apoptosis modulating protein; cell death; gene therapy;
KW Huntington's disease; neurotropic; anticonvulsant; cytostatic; mouse.
XX
XX Mus sp.
OS
XX
XX WO9960986-A2.
PN
XX
PD 02-DEC-1999.
XX
XX 27-MAY-1999; 99WO-US11743.
PF
XX 27-MAY-1998; 98US-0085199.
PR
XX (UYBR-) UNIV BRITISH COLUMBIA.
PA (MERI) MERCK FROSST CANADA INC.
XX

XX
PI Kalchman M, Hayden MR, Hackam A, Chopra V, Nicholson DW;
Vallaincourt JP, Rasper DM;
XX WPI: 2000-097055/08.
DR N-PSDB; T58748.
DR
XX Novel proteins useful for treating Huntington's disease by gene therapy
PT techniques, and cancers -
PT
XX Disclosure: Page 68-71; 91pp; English.
PS
XX The invention relates to Huntingtin Interacting Protein (HIP), that
CC includes a death effector domain (DED), suggesting apoptotic function.
CC Proteins with DED (referred as HIP-apoptosis modulating proteins) are
CC useful for inducing apoptotic death in cells. The HIP is a logical target
CC for therapy in Huntington's disease since it has a differential apoptotic
CC activity, modulated by interaction with Huntingtin having normal and
CC expanded repeats. HIP is also used as a therapeutic agent to introduce
CC apoptosis in cancer cells. Increasing expression of normal (non-expanded)
CC Huntingtin or the HIP-apoptotic modulator-binding portion, a modified
CC HIP-apoptotic modulator in which the DED has been deleted, is helpful for
CC treating Huntington's disease. The present sequence represents the mouse
CC mH1P1.
XX
XX Sequence 756 AA;
SQ

Query Match 61.0%; Score 36; DB 21; Length 756;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Oy 2 KLPRSPADSLT 12
:|:|||||
Db 431 rappepads1t 441

RESULT 13
W18030
ID W18030 standard; Protein; 914 AA.
XX
XX W18030;
AC
XX 03-SEP-1997 (first entry)
DT
XX Huntingtin interacting protein HIP1.
DE
XX Huntingtin interacting protein; HIP1; Huntington's disease;
KW therapy; diagnosis.
KW
XX Homo sapiens.
OS
XX WO9718825-A1.
PN
XX 29-MAY-1997.
PD
XX 15-NOV-1996; 96WO-US18370.
PF
XX 17-NOV-1995; 95US-0006882.
PR
XX (UYBR-) UNIV BRITISH COLUMBIA.
PA
XX Hayden MR, Kalchman M;
PI
XX WPI; 1997-297870/27.
DR N-PSDB; T67187.
DR
XX cDNA encoding Huntington's disease interacting protein - useful to
PT ameliorate effects of disease in patient expressing Huntingtin
PT protein with expanded CAG repeat region
XX
XX Claim 4; Page 28-31; 39pp; English.
PS
XX

CC IV-C) encoded by the DNA contained within the vector pc3-hpDE IV-C. The
 CC full length wild type hpDE IV-C coding sequence (see T90951) was isolated
 CC from a human testis cDNA library. Cyclic phosphodiesterase enzymes (PDEs)
 CC are a family of enzymes that catalyse the degradation of cyclic
 CC nucleotides. Cyclic nucleotides are important intracellular second
 CC messengers. This sequence can be used to produce the recombinant
 CC isoenzyme, which may be useful in PDE IV activity assays. The recombinant
 CC isoenzyme may also be used in screening assays for drugs that may be
 CC improved therapeutics in the areas of asthma and inflammation. Primers
 CC determined from this sequence, that are specific for hpDE IV-C, can be
 CC used in a RT-PCR amplification, in an assay for detecting hpDE IV-C in
 CC human cells.

XX
 SQ Sequence 506 AA;

Query Match 64.4%; Score 38; DB 18; Length 506;
 Best Local Similarity 66.7%; Pred. No. 46;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
 :|:||||:|
 Db 430 skiprpsd 438

RESULT 9

ID W34312 standard; protein; 519 AA.

XX W34312;

XX 26-FEB-1998 (first entry)

XX Human phosphodiesterase IV-C isoenzyme #3.

XX Human; cyclic nucleotide phosphodiesterase IV-C; isoenzyme; therapy;
 KW asthma; inflammation; hpDE IV-C.

XX Homo sapiens.

XX US5686286-A.

XX 11-NOV-1997.

XX 25-AUG-1993; 93US-0112815.

XX 05-AUG-1994; 94US-0286856.

PR 25-AUG-1993; 93US-0112815.

PR 07-JUN-1995; 95US-0472831.

XX (PFTZ) PFIZER INC.

XX Fisher DA;

XX WPI; 1997-558143/51.

DR N-PSDB; T93221.

XX Human phosphodiesterase IV isoenzyme hpDE IV-C - used to identify
 PT PDE inhibitors that may be used for treating asthma and inflammation

XX Disclosure; Fig 2; 13pp; English.

XX W34310-W34312 represent hpDE IV-C (human cyclic nucleotide
 CC phosphodiesterase IV-C) isoenzymes. The DNA encoding these sequences was
 CC isolated from a human testis cDNA library. The DNA encoding these
 CC sequences expressed by a host cell, can be used to determine the
 CC sequences of hpDE IV-C specific primers. The primers can be used for
 CC detecting the presence of hpDE IV-C in human cells. The host cell line
 CC can be used to identify compounds or other substances that inhibit or
 CC modify the activity of hpDE IV-C. The screening can identify drugs that
 CC may be improved therapeutics for treating asthma and inflammation.

XX Sequence 519 AA;

Query Match 64.4%; Score 38; DB 18; Length 519;
 Best Local Similarity 66.7%; Pred. No. 48;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
 :|:||||:|
 Db 443 skiprpsd 451

RESULT 10

ID R97867 standard; Protein; 606 AA.

XX R97867;

XX 13-FEB-1997 (first entry)

XX Human phosphodiesterase type IVC.

XX Phosphodiesterase; screening; identification; inhibitor;
 KW inhibition; PDE; treatment; prophylaxis; inflammatory disease;
 KW inflamed lung; asthma.

XX Homo sapiens.

XX WO9620281-A1.

XX 04-JUL-1996.

XX 21-DEC-1995; 95WO-GB03006.

XX 26-JUN-1995; 95GB-0012996.

PR 23-DEC-1994; 94GB-0026227.

XX (CLLT) CELLTech THERAPEUTICS LTD.

XX Lumb SM, Owens RJ, Perry MJ;

XX WPI; 1996-321854/32.

DR N-PSDB; T34621.

XX Human phosphodiesterase type IVC and selective inhibitors - used in
 PT the treatment of inflammatory disease, esp. asthma.

XX Claim 1; Figure 1; 50pp; English.

XX Recombinant phosphodiesterase (PDE) type IVC may be used to screen
 CC for inhibitors of PDE IVC. The inhibitors may be used in
 CC pharmaceutical for the treatment and prophylaxis of inflammatory
 CC diseases, especially inflamed lung associated with asthma.

XX Sequence 606 AA;

Query Match 64.4%; Score 38; DB 17; Length 606;
 Best Local Similarity 66.7%; Pred. No. 56;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
 :|:||||:|
 Db 530 skiprpsd 538

RESULT 11

ID R99742 standard; Protein; 606 AA.

XX R99742;

XX 13-FEB-1997 (first entry)

XX


```

XX OS Homo sapiens.
XX PN US5672509-A.
XX PD 30-SEP-1997.
XX PF 25-AUG-1993; 93US-0112815.
XX PR 05-AUG-1994; 94US-0286856.
XX PR 25-AUG-1993; 93US-0112815.
XX PA (PFIZ ) PFIZER INC.
XX PI Fisher DA;
XX DR WPI; 1997-488862/45.
XX DR N-PSDB; T90951.
XX PT DNA encoding human phosphodiesterase IV isoenzyme - useful for
XX PT producing recombinant isoenzyme, for screening for therapeutics for
XX PT asthma and inflammation
XX PS Disclosure; Column 11-14; 15pp; English.
XX CC This sequence represents the full length sequence for human
XX CC phosphodiesterase IV isoenzyme (hPDE IV-C). The DNA encoding this
XX CC sequence was isolated from a human testis cDNA library. Cyclic
XX CC phosphodiesterase enzymes (PDEs) are a family of enzymes that catalyze
XX CC the degradation of cyclic nucleotides. Cyclic nucleotides are important
XX CC intracellular second messengers. This sequence can be used to produce the
XX CC recombinant isoenzyme, which may be useful in PDE IV activity assays. The
XX CC recombinant isoenzyme may also be used in screening assays for drugs that
XX CC may be improved therapeutics in the areas of asthma and inflammation.
XX CC Primers determined from this sequence, that are specific for hPDE IV-C,
XX CC can be used in a RT-PCR amplification, in an assay for detecting hPDE
XX CC IV-C in human cells.
XX SQ Sequence 481 AA;

Query Match 64.4%; Score 38; DB 18; Length 481;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
Db 405 skiprpsd 413
:|:||||:|

RESULT 7
W34311
ID W34311 standard; protein; 506 AA.
XX AC W34311;
XX DT 26-FEB-1998 (first entry)
XX DE Human phosphodiesterase IV-C isoenzyme #2.
XX KW Human; cyclic nucleotide phosphodiesterase IV-C; isoenzyme; therapy;
XX KW asthma; inflammation; hPDE IV-C.
XX OS Homo sapiens.
XX PN US5686286-A.
XX PD 11-NOV-1997.
XX PF 25-AUG-1993; 93US-0112815.
XX PR 05-AUG-1994; 94US-0286856.
XX PR 25-AUG-1993; 93US-0112815.

```

```

PR 07-JUN-1995; 95US-0472831.
XX (PFIZ ) PFIZER INC.
XX PI Fisher DA;
XX DR WPI; 1997-558143/51.
XX PT Human phosphodiesterase IV isoenzyme hPDE IV-C - used to identify
XX PT PDE inhibitors that may be used for treating asthma and inflammation
XX PS Claim 1; Column 15-18; 13pp; English.
XX CC W34310-W34312 represent hPDE IV-C (human cyclic nucleotide
XX CC phosphodiesterase IV-C) isoenzymes. The DNA encoding these sequences was
XX CC isolated from a human testis cDNA library. The DNA encoding these
XX CC sequences expressed by a host cell, can be used to determine the
XX CC sequences of hPDE IV-C specific primers. The primers can be used for
XX CC detecting the presence of hPDE IV-C in human cells. The host cell line
XX CC can be used to identify compounds or other substances that inhibit or
XX CC modify the activity of hPDE IV-C. The screening can identify drugs that
XX CC may be improved therapeutics for treating asthma and inflammation.
XX SQ Sequence 506 AA;

Query Match 64.4%; Score 38; DB 18; Length 506;
Best Local Similarity 66.7%; Pred. No. 46;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLPRSPAD 9
Db 430 skiprpsd 438
:|:||||:|

RESULT 8
W30205
ID W30205 standard; Protein; 506 AA.
XX AC W30205;
XX DT 19-JAN-1998 (first entry)
XX DE Vector pc3-hPDE IV-C phosphodiesterase IV isoenzyme.
XX KW Phosphodiesterase IV isoenzyme; hPDE IV-C; human; PDE; enzyme; therapy;
XX KW cyclic nucleotide degradation; intracellular; second messenger; asthma;
XX KW inflammation.
XX OS Homo sapiens.
XX PN US5672509-A.
XX PD 30-SEP-1997.
XX PF 25-AUG-1993; 93US-0112815.
XX PR 05-AUG-1994; 94US-0286856.
XX PR 25-AUG-1993; 93US-0112815.
XX PA (PFIZ ) PFIZER INC.
XX PI Fisher DA;
XX DR WPI; 1997-488862/45.
XX PT DNA encoding human phosphodiesterase IV isoenzyme - useful for
XX PT producing recombinant isoenzyme, for screening for therapeutics for
XX PT asthma and inflammation
XX PS Disclosure; Column 13-18; 15pp; English.
XX CC This sequence represents the human phosphodiesterase IV isoenzyme (hPDE

```

QY 1 AKLPRSPAD 9
:|||||:
Db 146 skiprpsd 154

RESULT 4

Y49820
ID Y49820 standard; Protein; 253 AA.

AC Y49820;

XX 19-JAN-2000 (first entry)

DE Human temporal lobe phosphodiesterase pPDE21 protein.

KW Phosphodiesterase; duncce-like phosphodiesterase; PDE; DPD; cAMP;
KW RAS-related protein; immunoreactive; detection; genetic defect;
KW bronchodilation; increased myocardial contractility;
KW anti-inflammation.

OS Homo sapiens.

XX US977305-A.

PN 02-NOV-1999.

XX 07-JUN-1995; 95US-0474379.

PR 01-MAR-1994; 94US-0206188.

PR 20-APR-1990; 90US-0511715.

PR 19-APR-1991; 91US-0688352.

PA (COLD-) COLD SPRING HARBOR LAB.

XX Colicelli JJ, Wigler MH;

PI WPI; 1999-619709/53.

DR N-PSDB; 232254.

XX New isolated RAS-related polypeptides and mammalian cyclic nucleotide
phosphodiesterases, used for screening for agents which can modify
complement or suppress genetic defects -

PS Claim 2; Column 151-152; 145pp; English.

CC The present invention describes new isolated RAS-related polypeptides
and mammalian cyclic nucleotide phosphodiesterases (PDEs). RAS-related
polypeptides are capable of complementing a defective RAS function in
yeast. The products can be used for screening for agents which can
modify, complement or suppress a genetic defect in a biochemical
pathway in which CAMP participates, or in a biochemical pathway which
is controlled, directly or indirectly, by a RAS protein and other
proteins affecting cell growth and maintenance. Developing agents that
will selectively act upon PDEs is directed toward reproducing the
desirable effects of cyclic nucleotides, e.g. bronchodilation,
increased myocardial contractility, anti-inflammation, yet without
causing the undesirable effects, e.g. increased heart rate or enhanced
lipolysis. The products can also be used for therapeutic, diagnostic
and prognostic uses. 232229 to 232285, and Y49803 to Y49830, represent
sequences used in the exemplification of the present invention.

SQ Sequence 253 AA;

Query Match 64.4%; Score 38; DB 20; Length 253;

Best Local Similarity 66.7%; Pred. No. 22;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLPRSPAD 9

:|||||:

Db 177 skiprpsd 185

RESULT 5

W34310

ID W34310 standard; protein; 481 AA.

XX W34310;

XX 26-FEB-1998 (first entry)

XX Human phosphodiesterase IV-C isoenzyme #1.

XX Human; cyclic nucleotide phosphodiesterase IV-C; isoenzyme; therapy;
asthma; inflammation; hPDE IV-C.

OS Homo sapiens.

PN US5686286-A.

XX 11-NOV-1997.

XX 25-AUG-1993; 93US-0112815.

XX 05-AUG-1994; 94US-0286856.

XX 25-AUG-1993; 93US-0112815.

XX 07-JUN-1995; 95US-0472831.

XX (PFIZ) PFIZER INC.

XX Fisher DA;

XX WPI; 1997-558143/51.

XX N-PSDB; T93221.

XX Human phosphodiesterase IV isoenzyme hPDE IV-C - used to identify
PDE inhibitors that may be used for treating asthma and inflammation

XX Claim 1; Column 11-14; 13pp; English.

CC W34310-W34312 represent hPDE IV-C (human cyclic nucleotide
phosphodiesterase IV-C) isoenzymes. The DNA encoding these sequences was
isolated from a human testis cDNA library. The DNA encoding these
sequences expressed by a host cell, can be used to determine the
sequences of hPDE IV-C specific primers. The primers can be used for
detecting the presence of hPDE IV-C in human cells. The host cell line
can be used to identify compounds or other substances that inhibit or
modify the activity of hPDE IV-C. The screening can identify drugs that
may be improved therapeutics for treating asthma and inflammation.

SQ Sequence 481 AA;

Query Match 64.4%; Score 38; DB 18; Length 481;

Best Local Similarity 66.7%; Pred. No. 44;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLPRSPAD 9

:|||||:

Db 405 skiprpsd 413

RESULT 6

W30204

ID W30204 standard; Protein; 481 AA.

XX W30204;

XX 19-JAN-1998 (first entry)

XX Phosphodiesterase IV isoenzyme.

XX Phosphodiesterase IV isoenzyme; hPDE IV-C; human; PDE; enzyme; therapy;
cyclic nucleotide degradation; intracellular; second messenger; asthma;
inflammation.

PT hyper-calcaemia, hypo- or hyper-phosphatemia, kidney stones, etc
PS Claim 9; Fig 1A-E; 62pp; English.
CC A novel 7-transmembrane receptor (W12695) has been identified as a
CC human G-protein parathyroid hormone (PTH) receptor, designated
CC HLTG74. It shows 48.2% homology to the human PTH receptor. Its
CC amino acid sequence was deduced from a cDNA clone (T59619) isolated
CC from a human T cell lymphoma tissue cDNA library. Recombinant
CC HLTG74 can be produced in transfected host cells and used to
CC screen for (ant)agonist cpds. Agonists may be used to prevent or
CC treat e.g. hypocalcaemia, hyperphosphataemia, hypoparathyroidism
CC and chronic tetany by stimulating an increase in serum calcium
CC levels. Antagonists can be used to inhibit the receptor e.g. for
CC the treatment of osteoporosis, hypercalcaemia, hypoparathyroidism,
CC hypophosphataemia, kidney stone, nephrolithiasis.
XX SQ Sequence 541 AA;

Query Match 100.0%; Score 59; DB 18; Length 541;
Best Local Similarity 100.0%; Pred. No. 0.0089;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLPRSPADSLT 12
Db 476 akprspadsit 487
|||||||

RESULT 2
R14852
ID R14852 standard; Protein; 222 AA.
XX AC R14852;
XX 30-JAN-1992 (first entry)
XX Protein associated with biochemical pathway involving cAMP.
DE RAS; oncogene; cancer; cyclic nucleotide phosphodiesterase; ss.
XX Homo sapiens.
XX W09116457-A.
XX 31-OCT-1991.
XX 19-APR-1991; 91WO-US02714.
XX 20-APR-1990; 90US-0511715.
XX (COLD-) COLD SPRING HARBOR.
XX Wigler MH, Colicelli JJ;
XX WPI; 1991-339841/46.
XX N-PSDB; Q14643.
XX Complementary screening for genes and prods. - e.g. RAS protein
PT and cAMP, that modify, complement or suppress genetic defect and
PT correct associated phenotypic alteration
XX Disclosure; Page 132; 169pp; English.
XX In the specification this sequence is given the SEQ ID NO. 43
CC which is described as the nucleotide (1) sequence of the insert
CC from plasmid pPDE21 which encodes a member of the cAMP-specific
CC phosphodiesterase family.
XX SQ Sequence 222 AA;

Query Match 64.4%; Score 38; DB 12; Length 222;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 66.7%; Pred. No. 19;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLPRSPAD 9
Db 146 skiprpsd 154
:|:||||:|

RESULT 3
W00100
ID W00100 standard; Protein; 222 AA.
XX AC W00100;
XX 10-OCT-1996 (first entry)
XX Family IV phosphodiesterase encoded by plasmid pPDE21 insert.
XX Human; glioblastoma cell; plasmid; mammalian; complementation; pPDE2RR;
KW probe; yeast; pPDE1; pPDE10X inv; temporal lobe; cDNA library; PRATPDP;
KW pTM72; pGB14; pGB18ARR; pTM3; pJC44X; pGB25; phosphodiesterase family IV;
KW pPDE18; pPDE21.
XX Homo sapiens.
OS US5527896-A.
XX 18-JUN-1996.
XX 20-APR-1990; 90US-0511715.
XX 19-APR-1991; 91US-0688352.
XX 20-APR-1990; 90US-0511715.
XX (COLD-) COLD SPRING HARBOR LAB.
XX Colicelli JJ, Wigler MH;
XX WPI; 1996-299902/30.
XX N-PSDB; T34384.
XX DNA mols. isolated from human glioblastoma cells - encode
PT RAS-related or cyclic nucleotide phosphodiesterase proteins
XX Claim 5; Column 139-142; 101pp; English.
XX The sequences given in W00095-100 are encoded by plasmid fragments
CC which were isolated by hybridisation with mammalian genes cloned by
CC complementation. The cDNA sequences were isolated using probes
CC derived from the sequences given in T34366-76 which were cloned via
CC complementation in yeast. Plasmids pPDE1, pPDE10X inv and pPDE2RR were
CC isolated by low stringency hybridisation screens of a human temporal
CC lobe cDNA library using the PRATPDP insert as a probe. Comparison of
CC the nucleotide sequences given in T34377-79 indicated that the inserts
CC are representatives of the same genetic locus as the insert in pTM72.
CC Plasmids pGB14 and pGB18ARR were obtained in the same manner. DNA
CC sequence analysis revealed that they are representatives of the same
CC genetic locus as the inserts in pTM3 and pJC44X. Plasmid pGB25 was also
CC obtained at low stringency hybridisation using the PRATPDP insert as a
CC probe. Judged by its nucleotide and deduced amino acid sequence it
CC represents a novel member of the phosphodiesterase family IV. The cDNA
CC insert of pGB25 was used as a probe to obtain pPDE18 and pPDE21. The
CC cDNA of pPDE18 represents the same locus as that of pGB25 and contains
CC more sequence information than the pGB25 cDNA. The pPDE21 insert
CC represents a fourth member of phosphodiesterase family IV. The
CC assignment to family IV is based solely on sequence relationships.
XX SQ Sequence 222 AA;

Query Match 64.4%; Score 38; DB 17; Length 222;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:51:49 ; Search time 138.73 Seconds
(without alignments)
2.958 Million cell updates/sec

Title: us-09-236-468a-2_COPY_476_487
Perfect score: 59
Sequence: 1 AKLPRSPADSLT 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues
Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*

1: /SIDSL1/gcgdata/geneseq/geneseq/AA1980.DAT:*

2: /SIDSL1/gcgdata/geneseq/geneseq/AA1981.DAT:*

3: /SIDSL1/gcgdata/geneseq/geneseq/AA1982.DAT:*

4: /SIDSL1/gcgdata/geneseq/geneseq/AA1983.DAT:*

5: /SIDSL1/gcgdata/geneseq/geneseq/AA1984.DAT:*

6: /SIDSL1/gcgdata/geneseq/geneseq/AA1985.DAT:*

7: /SIDSL1/gcgdata/geneseq/geneseq/AA1986.DAT:*

8: /SIDSL1/gcgdata/geneseq/geneseq/AA1987.DAT:*

9: /SIDSL1/gcgdata/geneseq/geneseq/AA1988.DAT:*

10: /SIDSL1/gcgdata/geneseq/geneseq/AA1989.DAT:*

11: /SIDSL1/gcgdata/geneseq/geneseq/AA1990.DAT:*

12: /SIDSL1/gcgdata/geneseq/geneseq/AA1991.DAT:*

13: /SIDSL1/gcgdata/geneseq/geneseq/AA1992.DAT:*

14: /SIDSL1/gcgdata/geneseq/geneseq/AA1993.DAT:*

15: /SIDSL1/gcgdata/geneseq/geneseq/AA1994.DAT:*

16: /SIDSL1/gcgdata/geneseq/geneseq/AA1995.DAT:*

17: /SIDSL1/gcgdata/geneseq/geneseq/AA1996.DAT:*

18: /SIDSL1/gcgdata/geneseq/geneseq/AA1997.DAT:*

19: /SIDSL1/gcgdata/geneseq/geneseq/AA1998.DAT:*

20: /SIDSL1/gcgdata/geneseq/geneseq/AA1999.DAT:*

21: /SIDSL1/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	541	W12695	G-protein parathyr
2	38	64.4	222	R14852	Protein associated
3	38	64.4	222	W00100	Family IV phospho
4	38	64.4	253	W49820	Human temporal lob
5	38	64.4	481	W34310	Human phosphodiesterase
6	38	64.4	481	W30204	Phosphodiesterase
7	38	64.4	506	W34311	Human phosphodiesterase
8	38	64.4	506	W30205	Vector pc3-hPDE IV
9	38	64.4	519	W34312	Human phosphodiesterase
10	38	64.4	606	W34312	Human phosphodiesterase
11	38	64.4	606	W34312	Human phosphodiesterase
12	36	61.0	756	Y59272	Mouse huntingtin-i

13	36	61.0	914	18	W18030	Huntingtin interac
14	36	61.0	914	21	Y59269	Human huntingtin-i
15	36	61.0	1090	21	Y59270	Human huntingtin-i
16	36	61.0	1138	21	Y83222	CAP6 polypeptide (
17	35	59.3	437	21	Y44847	Glycine max 3-dehy
18	34	57.6	108	20	Y12491	Human 5' EST secre
19	34	57.6	290	20	Y34851	Chlamydia pneumoni
20	34	57.6	436	21	Y44846	Oryza sativa 3-deh
21	34	57.6	438	21	Y44845	zea mays 3-dehydro
22	33	55.9	41	16	R71941	Alsef protein. Ho
23	33	55.9	94	21	Y50793	Retrotransposon Ta
24	33	55.9	265	16	R81423	Hepatitis GB virus
25	33	55.9	287	20	Y38607	Neisseria meningit
26	33	55.9	287	20	Y38609	Neisseria meningit
27	33	55.9	287	21	Y74901	Neisseria meningit
28	33	55.9	288	20	Y38610	Neisseria meningit
29	33	55.9	288	20	Y38611	Neisseria gonorrhoe
30	33	55.9	288	21	Y74899	Neisseria gonorrhoe
31	33	55.9	371	11	R06000	Neisseria gonorrhoe
32	33	55.9	386	20	Y16100	Bacteriophage lamb
33	33	55.9	403	20	W97719	Acetobacter xylinu
34	33	55.9	404	20	Y23858	Staphylococcus aur
35	33	55.9	427	19	W61355	A Trichomonas vagi
36	33	55.9	480	14	R33842	Human hypothalamic
37	33	55.9	480	14	R33841	Streptolysin O var
38	33	55.9	571	12	R10376	Soluble streptolys
39	33	55.9	855	19	W79529	Streptolysin O der
40	33	55.9	855	21	Y32214	Mouse CLOCK protei
41	33	55.9	1045	16	R66221	Mouse CLOCK protei
42	33	55.9	1050	20	W67642	Sucrose-phosphate-
43	33	55.9	1051	20	W67641	A serine/threonine
44	33	55.9	1182	20	Y33496	A serine/threonine
45	33	55.9	1853	21	Y53668	Human SCA6 protein
						Protein 608 sequen

ALIGNMENTS

RESULT 1	
W12695	
ID	W12695 standard; Protein; 541 AA.
XX	
AC	W12695:
XX	
DT	31-MAY-1997 (first entry)
XX	
DE	G-protein parathyroid hormone receptor HLTG74.
XX	
KW	G-protein parathyroid hormone receptor; HLTG74; parathormone; PTH;
KW	calcium; signal transduction; agonist; antagonist; hypocalcaemia;
KW	hyperphosphataemia; hypoparathyroidism; chronic tetany;
KW	osteoporosis; hypercalcaemia; hypoparathyroidism; hypophosphataemia;
KW	kidney stone; nephrolithiasis; therapy; diagnosis.
OS	Homo sapiens.
XX	
PN	W09639433-A1.
XX	
PD	12-DEC-1996.
XX	
PF	05-JUN-1995; 95WO-US07085.
XX	
PR	05-JUN-1995; 95WO-US07085.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Li Y, Rosen CA, Ruben SM, Soppet DR;
XX	
DR	WPI: 1997-043068/04.
XX	
DR	N-FSDB: Y59619.
XX	
PT	Human G-protein parathyroid hormone receptor, HLTG74 - used to identify (ant)agonists, used in the treatment of hypo- or

DR PFAM; PF00168; C2; 1.
 DR PFAM; PF00397; WW; 3.
 DR PFAM; PF00632; HECT; 1.
 DR PRINTS; PR00403; WWDOMAIN.
 DR PROSITE; PS00499; C2_DOMAIN_1; UNKNOWN_1.
 DR PROSITE; PS01159; WW_DOMAIN_1; 3.
 DR PROSITE; PS50004; C2_DOMAIN_2; 1.
 DR PROSITE; PS50020; WW_DOMAIN_2; 3.
 KW Hypothetical protein; Ubiquitin conjugation; Ligase; Repeat.
 FT DOMAIN 15 30
 FT DOMAIN 149 157 POLY-THR.
 FT DOMAIN 242 267 WW DOMAIN.
 FT DOMAIN 312 337 WW DOMAIN.
 FT DOMAIN 370 395 WW DOMAIN.
 FT DOMAIN 482 786 HECT DOMAIN.
 FT BINDING 754 754 UBIQUITIN (BY SIMILARITY).
 SQ SEQUENCE 786 AA; 89259 MW; 5787A859F5497B9A CRC64;

Query Match 79.5%; Score 31; DB 3; Length 786;
 Best Local Similarity 75.0%; Pred. No. 50;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
 Db 275 HNTSSDSQ 282

Search completed: November 8, 2000, 08:56:18
 Job time: 529 sec

```
OC Rhabdittidae; Peloderinae; Caenorhabdittis.
RN [1]
RP SEQUENCE FROM N.A.
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RL Lloyd C.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBDJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL Nature 368:32-38(1994).";
DR EMBL; Z68159; CAA92286.1; -.
SQ SEQUENCE 604 AA; 70472 MW; 609CF44D2854E624 CRC64;

Query Match 79.5%; Score 31; DB 5; Length 604;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSTSSQS 7
   |||||
Db 7 HSTSSQS 13

RESULT 13
O45894 PRELIMINARY; PRT; 631 AA.
AC O45894;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE W06D12.5 PROTEIN.
GN W06D12.5.
OS Caenorhabdittis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdittidae; Rhabditoidea;
OC Rhabdittidae; Peloderinae; Caenorhabdittis.
RN [1]
RP SEQUENCE FROM N.A.
RA Basham V.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBDJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL Nature 368:32-38(1994).";
DR EMBL; Z82073; CAB04926.1; -.
DR INTERPRO; IPR001622; -.
SQ SEQUENCE 631 AA; 71794 MW; C4E56637B0614CFC CRC64;

Query Match 79.5%; Score 31; DB 5; Length 631;
Best Local Similarity 75.0%; Pred. No. 41;

OC Rhabdittidae; Peloderinae; Caenorhabdittis.
RN [1]
RP SEQUENCE FROM N.A.
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RL Lloyd C.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBDJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL Nature 368:32-38(1994).";
DR EMBL; Z68159; CAA92286.1; -.
DR INTERPRO; IPR001622; -.
SQ SEQUENCE 604 AA; 70472 MW; 609CF44D2854E624 CRC64;

Query Match 79.5%; Score 31; DB 3; Length 650;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 STSSQS 8
   |||||
Db 404 STSSQS 410

RESULT 15
O14326 PRELIMINARY; PRT; 786 AA.
AC O14326;
DT 01-JAN-1999 (TREMBlrel. 09, Created)
DT 01-JAN-1999 (TREMBlrel. 09, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE PUTATIVE UBIQUITIN--PROTEIN LIGASE C16E9.11C (EC 6.3.2.-).
GN SPEC16E9.11C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
OC Schizosaccharomycetales; Schizosaccharomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA Volckaert G., Wood V., Rajandream M.A., Barrell B.G.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBDJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE = AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION.
CC -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
CC -!- SIMILARITY: CONTAINS 3 WW DOMAINS.
CC -!- SIMILARITY: CONTAINS AN HECT-TYPE E3 UBIQUITIN-PROTEIN LIGASE
CC DOMAIN.
CC -!- SIMILARITY: STRONG, TO S.POMBE PUB1 AND YEAST RSP5.
DR EMBL; Z99759; CAB16903.1; -.
DR HSSP; O13526; 1PIN.
DR INTERPRO; IPR000008; -.
DR INTERPRO; IPR000569; -.
DR INTERPRO; IPR001202; -.
DR INTERPRO; IPR002349; -.

```

OS Lentinula edodes (Shiitake mushroom) (Lentinus edodes).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Agaricales;
OC Tricholomataceae; Lentinula.

[1]

RN SEQUENCE FROM N.A.
RP STRAIN-MSLE1;
RC SHISHIDO K.;
RA "mfa homolog 2, encoding cell-adhesion protein, from the
RT basidiomycete Lentinus edodes strain MSLE1.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB028177; BAA78613.1; -;
FT NON_TER 1 1
FT 352 352
SQ SEQUENCE 352 AA; 39001 MW; BBFE4018BE43CAA8 CRC64;

Query Match 79.5%; Score 31; DB 3; Length 352;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HSTSSSQ 8
|||:|
Db 165 HSTASSQ 172

RESULT 9

ID Q9UQX4 PRELIMINARY; PRT; 352 AA.
AC Q9UQX4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE CELL-ADHESION PROTEIN (FRAGMENT).
GN MFBA.
OS Lentinula edodes (Shiitake mushroom) (Lentinus edodes).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Agaricales;
OC Tricholomataceae; Lentinula.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-MSLE1, AND FMC2;
RC SHISHIDO K.;
RA "mfa homolog1, encoding cell-adhesion protein, from the basidiomycete
RT Lentinus edodes strain MSLE1.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB028176; BAA78612.1; -;
DR EMBL: AB028175; BAA78611.1; -;
FT NON_TER 1 1
FT 352 352
SQ SEQUENCE 352 AA; 38998 MW; DD0398489D02781B CRC64;

Query Match 79.5%; Score 31; DB 3; Length 352;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HSTSSSQ 8
|||:|
Db 165 HSTASSQ 172

RESULT 10

ID Q9SUD5 PRELIMINARY; PRT; 359 AA.
AC Q9SUD5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE MEDICAGO NODULIN N21-LIKE PROTEIN.
GN T13J8.150.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Rosidae; eurosids II; Brassicales;
OC Brassicaceae; Arabidopsis.

[1]
RN SEQUENCE FROM N.A.
RP Bevan M., Pohl T., Weizenegger T., Hoheisel J., Mewes H.W.,
RA Mayer K.F.X., Schueller C.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]

RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL035524; CAB36773.1; -;
SQ SEQUENCE 359 AA; 39347 MW; BDDER39DB3E1C4C9 CRC64;

Query Match 79.5%; Score 31; DB 10; Length 359;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSTSSSQ 7
|||:|
Db 330 HNTSSQ 336

RESULT 11

ID Q99113 PRELIMINARY; PRT; 405 AA.
AC Q99113;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE HOMEODOMAIN PROTEIN BW3 (FRAGMENT).
GN BW3.
OS Ustilago maydis (Smut fungus).
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-RK32;
RC MEDLINE; 92154679.
RA Gillissen B., Bergemann J., Sandmann C., Schroeder B., Boelker M.,
RA Kahmann R.;
RA "A two-component regulatory system for self/non-self recognition in
RT Ustilago maydis.";
RL Cell 68:647-657(1992).
CC -|- SUBCELLULAR LOCATION: NUCLEAR.
DR EMBL: M84180; AAA34222.1; -;
DR INTERPRO: IPR001356; -;
DR PFAM: PF00046; homeobox; 1.
DR PROSITE: PS00071; HOMEBOX_2; 1.
KW Transcription regulation; Homeobox; DNA-binding; Nuclear protein.
FT NON_TER 405 405
SQ SEQUENCE 405 AA; 45316 MW; 96C7255922AF904D CRC64;

Query Match 79.5%; Score 31; DB 3; Length 405;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HSTSSSQ 8
|||:|
Db 292 HSTSSSE 299

RESULT 12

ID Q18376 PRELIMINARY; PRT; 604 AA.
AC Q18376;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)
DE C33D9.8 PROTEIN.
GN C33D9.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

```

RESULT 5
Q9VX75 PRELIMINARY; PRT; 1464 AA.
AC Q9VX75;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE BAZ PROTEIN.
GN BAZ.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE; 20196006.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman J.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cusack B., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Flesher C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Stiden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003504; AAF48703.1; -.
DR FLYBASE; FBgn0000163; baz.
DR INTERPRO; IPR001478; -.
DR PFAM; PF00595; PDZ; 3.
SQ SEQUENCE 1464 AA; 157426 MW; 089B43434CA818C3 CRC64;

```

```

Query Match 82.18; Score 32; DB 5; Length 1464;
Best Local Similarity 87.5%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 6
Qy 1 HSTSSQSQ 8
Db 1320 HSTSSQSQ 1327

```

```

Q9UWE8 PRELIMINARY; PRT; 350 AA.
AC Q9UWE8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE CELL-ADHESION PROTEIN (FRAGMENT).
OS Lentinula edodes (Shiitake mushroom) (Lentinus edodes).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Agaricales;
OC Tricholomataceae; Lentinula.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSLE3;
RA Shishido K.;
RT "mfa homolog 2, encoding cell-adhesion protein, from the
RT basidiomycete Lentinus edodes strain MSLE3.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB028179; BAA78615.1; -.
FT NON_TER 1 350
FT NON_TER 350 350
SQ SEQUENCE 350 AA; 38813 MW; CDBF8258F301F158 CRC64;

Query Match 79.5%; Score 31; DB 3; Length 350;
Best Local Similarity 75.0%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
Db 165 HSTASSQ 172

RESULT 7
Q9Y742 PRELIMINARY; PRT; 352 AA.
AC Q9Y742;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE CELL-ADHESION PROTEIN (FRAGMENT).
OS Lentinula edodes (Shiitake mushroom) (Lentinus edodes).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Agaricales;
OC Tricholomataceae; Lentinula.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSLE3;
RA Shishido K.;
RT "mfa homolog 1, encoding cell-adhesion protein, from the
RT basidiomycete Lentinus edodes strain MSLE3.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB028178; BAA78614.1; -.
FT NON_TER 1 352
FT NON_TER 352 352
SQ SEQUENCE 352 AA; 38751 MW; D75992306014E4CC CRC64;

Query Match 79.5%; Score 31; DB 3; Length 352;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
Db 165 HSTASSQ 172

RESULT 8
Q9Y741 PRELIMINARY; PRT; 352 AA.
AC Q9Y741;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE CELL-ADHESION PROTEIN (FRAGMENT).

```



```

DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DE PARATHYROID HORMONE TYPE-2 RECEPTOR PRECURSOR.
OS Brachydanio retio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprinidae; Rasbora; Danio.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE: 99367425.
RA Rubin D.A., Hellman P., Zon L.I., Lobb C.J., Bergwitz C., Juppner H.;
RT "A G protein-coupled receptor from zebrafish is activated by human
RT parathyroid hormone and not by human or teleost parathyroid hormone-
RT related peptide. Implications for the evolutionary conservation of
RT calcium-regulating peptide hormones."
RL J. Biol. Chem. 274:23035-23042(1999).
DR EMBL: AF132082; AAD51908.1; -
DR INTERPRO: IPR000832; -
DR INTRPRO: IPR001879; -
DR PFAM: PF00002; 7tm_2; 1.
DR PRINTS: PS00249; GPCRSECRETIN.
DR PROSITE: PS00649; G_PROTEIN_RECPT_F2_1; 1.
DR PROSITE: PS00650; G_PROTEIN_RECPT_F2_2; 1.
KW Receptor.
SQ SEQUENCE 575 AA; 64244 MW; 888F1C4DDB3A14DC CRC64;

Query Match 82.1%; Score 32; DB 13; Length 575;
Best Local Similarity 87.5%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSTSSSQSQ 8
Db 484 NSTSSSQSQ 491
:|||||

RESULT 3
Q9VEP9 PRELIMINARY; PRT; 784 AA.
AC Q9VEP9
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DE CG16941 PROTEIN.
GN CG16941.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE: 20196006.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananthites P.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew K.Y., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.M., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de-Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

```

```

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacht J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003715; AAF55372.1; -
DR FLYBASE: FBgn0038464; CG16941.
DR INTERPRO: IPR000061; -
DR INTRPRO: IPR000626; -
DR PFAM: PF00240; ubiquitin; 1.
DR PFAM: PF01805; Surp; 2.
DR PRINTS: PR00348; UBIQUITIN.
DR PROSITE: PS00053; UBIQUITIN_2; 1.
DR PROSITE: PS00053; UBIQUITIN_2; 1.
SQ SEQUENCE 784 AA; 88076 MW; 9E6BCDF18E650A93 CRC64;

Query Match 82.1%; Score 32; DB 5; Length 784;
Best Local Similarity 87.5%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HSTSSSQSQ 8
Db 576 HSHSSSQSQ 583
:|||||

RESULT 4
O96782 PRELIMINARY; PRT; 1464 AA.
AC O96782
DT 01-MAY-1999 (TEMBLrel. 10, Created)
DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
DT 01-JUN-2000 (TEMBLrel. 14, Last annotation update)
DE BAZOOKA PROTEIN.
GN BAZOOKA.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
RN [1]
RP SEQUENCE FROM N.A.
RA Kuchinke U., Grawe F., Knust E.;
RT "Control of spindle orientation in Drosophila by the Par-3-related
RT PDZ-domain protein Bazooka."
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ130871; CAA10224.1; -
DR HSSP: P31016; IBE9.
DR INTRPRO: IPR001478; -
DR PFAM: PF00595; PDZ; 3.
DR PFAM: PF00595; PDZ; 3.
SQ SEQUENCE 1464 AA; 157398 MW; 77381A5FBCA4CC4B CRC64;

Query Match 82.1%; Score 32; DB 5; Length 1464;
Best Local Similarity 87.5%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HSTSSSQSQ 8
Db 1320 HSTSSSQSQ 1327
:|||||

```



```
FT ZN_FING 346 371 C2H2-TYPE.
-FT ZN_FING 377 402 C2H2-TYPE.
FT ZN_TER 1360 1360
SQ SEQUENCE 1360 AA; 149422 MW; 5A32B8086794BC2D CRC64;

Query Match 74.4%; Score 29; DB 1; Length 1360;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HSTSSQS 7
|:|:|:|:|
Db 1175 HSVSSQS 1181

RESULT 14
CCR4_RAT
ID CCR4_RAT STANDARD; PRT; 349 AA.
AC Q08565;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (CXCR-4) (SDF-1 RECEPTOR)
DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN) (LEUKOCYTE-DERIVED
DE SEVEN TRANSMEMBRANE DOMAIN RECEPTOR) (LESTR).
GN CXCR4 OR CMKAR4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
SEQUENCE FROM N.A.
RC STRAIN-WISTAR; TISSUE-SPLEEN;
RA Harrison J.K.; Saltafranca M.N.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases
CC -1- FUNCTION: RECEPTOR FOR THE C-X-C CHEMOKINE SDF-1. TRANSDUCES A
CC SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U90610; AAB50408.1; -
CC GCRDB; GCR_1401; -
CC INTERPRO; IPR000276; -
CC INTERPRO; IPR001277; -
CC PFAM; PF00001; 7tm.1; 1.
CC PRINTS; PR00237; GPCRHDOPSN.
CC PRINTS; PR00645; ICRIOHPHARN.
CC PROSITE; PS00237; G_PROTEIN_RECEPTOR; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 36
FT TRANSMEM 37 60 1 (POTENTIAL).
FT DOMAIN 61 76 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 77 96 2 (POTENTIAL).
FT DOMAIN 97 107 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 108 129 3 (POTENTIAL).
FT DOMAIN 130 151 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 152 172 4 (POTENTIAL).
FT DOMAIN 173 197 5 (POTENTIAL).
FT TRANSMEM 198 217 6 (POTENTIAL).
FT DOMAIN 218 237 7 (POTENTIAL).
FT TRANSMEM 238 258 8 (POTENTIAL).
FT DOMAIN 259 282 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 283 302 7 (POTENTIAL).
FT DOMAIN 303 349 CYTOPLASMIC (POTENTIAL).
FT DISULFID 106 183 BY SIMILARITY.
FT CARBOHYD 8 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
SQ SEQUENCE 349 AA; 39334 MW; 7E0789A605C60C09 CRC64;

Query Match 71.8%; Score 28; DB 1; Length 349;
Best Local Similarity 71.4%; Pred. No. 53;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSTSSQS 7
|:|:|:|:|
Db 137 HATNSQS 143

RESULT 15
FILA_HUMAN
ID FILA_HUMAN STANDARD; PRT; 416 AA.
AC P20930;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE FILAGGRIN PRECURSOR (FRAGMENT).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE; 89296901.
RA McKinley-Grant L.J.; Idler W.W.; Bernstein I.A.; Parry D.A.D.;
RA Cannizzaro L.; Croce C.M.; Huebner K.; Lessin S.R.; Steinert P.M.;
RT "Characterization of a cDNA clone encoding human filaggrin and
RT localization of the gene to chromosome region 1q21.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4848-4852(1989).
CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
CC DISULFIDE-BOND FORMATION AMONG THE INTERMEDIATE FILAMENTS DURING
CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC OF 324 AA, WHICH ARE NOT SEPARATED BY "LARGE LINKER". THE
CC PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES DURING TERMINAL
CC DIFFERENTIATION IT IS DEPHOSPHORYLATED & PROTEOLYTICALLY CLEAVED.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M24355; AAA52454.1; -
CC PIR; A32947; A32947.
CC MIM; 135940; -
CC KW Phosphorylation; Polyprotein; Developmental protein.
CC NON_TER 1
FT NON_TER 1
SQ SEQUENCE 416 AA; 44105 MW; DEEA3218BA043E32 CRC64;

Query Match 71.8%; Score 28; DB 1; Length 416;
Best Local Similarity 62.5%; Pred. No. 64;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HSTSSQS 8
|:|:|:|:|
Db 64 HSTSQEQ 71
```

Search completed: November 8, 2000, 09:03:55
Job time: 865 sec

```

RESULT 11
YA4C_SCHPO STANDARD; PRT; 596 AA.
AC Q09729;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 65.9 KDA PROTEIN C31A2.12 IN CHROMOSOME 1.
GN SPAC31A2.12.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
CC Schizosaccharomycetaceae; Schizosaccharomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Devlin K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO S.POMBE SPAC8A4.13C AND TO YEAST ROD1, YFR022W AND
CC YGR068C.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Z50113; CAA90470.1; -.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 194 211 POTENTIAL.
FT DOMAIN 532 537 POLY-SRP.
SQ SEQUENCE 596 AA; 65902 MW; C93B9FA55688E7B1 CRC64;

Query Match 74.4%; Score 29; DB 1; Length 596;
Best Local Similarity 85.7%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HSTSSQS 7
|||||
Db 531 HSTSSSS 537

RESULT 12
PEX1_PICPA STANDARD; PRT; 1157 AA.
AC P46463;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PEROXISOME BIOSYNTHESIS PROTEIN PASI (PEROXIN-1).
GN PEX1 OR PASI.
OS Pichia pastoris (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
CC Saccharomycetaceae; Pichia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=21-1;
RX MEDLINE; 95050987.
RA Heyman J.A., Monosov E., Subramani S.;
RL J. Cell Biol. 127:1259-1273(1994).
CC -!- FUNCTION: INVOLVED IN PEROXISOME BIOSYNTHESIS. MAY PLAY A DIRECT
CC OR INDIRECT ROLE IN DELIVERING MEMBRANE MATERIAL TO DEVELOPING
CC PEROXISOMES. IT MAY ALSO BE INVOLVED IN INTRACELLULAR MEMBRANE
CC MOVEMENT.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC -!- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; Z36987; CAA83450.1; -.
DR INTERPRO; IPR001939; -.
DR PFAM; PF00004; AAA; 2.
DR PROSITE; PS00674; AAA; 1.
KW Peroxisome; ATP-binding; Repeat.
FT NP_BIND 523 530 ATP (POTENTIAL).
FT NP_BIND 840 847 ATP (POTENTIAL).
SQ SEQUENCE 1157 AA; 126984 MW; B26F46F55A92DA3C CRC64;

Query Match 74.4%; Score 29; DB 1; Length 1157;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HSTSSQS 8
|||||
Db 597 HSDSSQR 604

RESULT 13
GLI1_XENLA STANDARD; PRT; 1360 AA.
AC Q91690;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ZINC FINGER PROTEIN GLI1 (GLI-1) (FRAGMENT).
GN GLI1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
CC Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97359968.
RA Lee J., Platt K.A., Censullo P., Ruiz I Altaba A.;
RT "Gli1 is a target of Sonic hedgehog that induces ventral neural tube
RL development.";
RN [2]
RP REVISIONS.
RA Lee J., Platt K.A., Censullo P., Ruiz I Altaba A.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: IMPLICATED IN THE TRANSDUCTION OF SHH SIGNAL (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: TO THE GLI-RELATED GROUP OF C2H2-TYPE ZINC-FINGERS
CC PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U57454; AAC24946.1; -.
DR HSPSP; P08151; 2GLI.
DR INTERPRO; IPR000822; -.
DR PFAM; PF00096; zf-C2H2; 5.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
KW Zinc-finger; Metal-binding; DNA-binding; Transcription regulation;
KW Nuclear protein.
FT DOMAIN 250 402 ZINC-FINGERS.
FT ZN_FING 250 275 ZINC-TYPE.
FT ZN_FING 316 340 C2H2-TYPE.

```

RX MEDLINE; 90191711.
RA Alcalay M., Antolini F., van de Ven W.J., Lanfranccone L.,
RA Grignani F., Pelicci P.G.;
RT "Characterization of human and mouse c-fes cDNA clones and
RT identification of the 5' end of the gene.";
RL Oncogene 5:267-275(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 86055727.
RA Roebroek A.J.M., Schalken J.A., Verbeek J.S., van den Ouweland A.M.W.,
RA Onnekink C., Bloemers H.P.J., van de Ven W.J.M.;
RT "The structure of the human c-fes/fps proto-oncogene.";
RL EMBL J. 4:2857-2903(1985)
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN. BELONGS TO THE FES/FPS SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X52192; CAA36438.1; -;
CC EMBL; X06292; CAA29619.1; -;
CC PIR; A24673; TVHUFF.
CC PIR; A60188; A60188.
CC HSP; P11362; IFGI.
CC MIM; 190030; -;
CC INTERPRO; IPR000719; -;
CC INTERPRO; IPR000980; -;
CC INTERPRO; IPR001060; -;
CC INTERPRO; IPR001245; -;
CC PFAM; PF00611; FCH; 1.
CC PFAM; PF00017; SH2; 1.
CC PFAM; PF00069; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC PRINTS; PR00401; SH2DOMAIN.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS50001; SH2; 1.
CC Transferase: Tyrosine-protein kinase; Proto-oncogene; ATP-binding;
KW Phosphorylation; SH2 domain.
FT DOMAIN 460 549 SH2.
FT DOMAIN 561 822 PROTEIN KINASE.
FT NP_BIND 567 575 ATP (BY SIMILARITY).
FT BINDING 590 590 ATP (BY SIMILARITY).
FT ACT_SITE 683 683 BY SIMILARITY.
FT MOD_RES 713 713 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 719 719 S -> L (IN REF. 2).
SQ SEQUENCE 822 AA; 93470 MW; ADA0B8F7D2666356 CRC64;

Query Match 76.9%; Score 30; DB 1; Length 822;
Best Local Similarity 75.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HSTSSQSQ 8
Db 407 HSTSSESQ 414
|||||

RESULT 10
DMP1_HUMAN STANDARD; PRT; 513 AA.
AC Q13316; O43265;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DENTIN MATRIX ACIDIC PHOSPHOPROTEIN 1 PRECURSOR (DENTIN MATRIX
DE PROTEIN-1) (DMP-1).
GN DMP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX TISSUE=MOLAR;
RX MEDLINE; 97321043.
RA Hirst K.L., Simmons D., Feng J., Aplin H., Dixon M.J., McDougall M.;
RT "Elucidation of the sequence and the genomic organization of the human
RT dentin matrix acidic phosphoprotein 1 (DMP1) gene: exclusion of the
RT locus from a causative role in the pathogenesis of dentinogenesis
RT imperfecta type II.";
RL Genomics 42:38-45(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA McDougall M., Juan X., Simmons D., Feng J.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 462-513 FROM N.A., AND GENE MAPPING.
RX MEDLINE; 96163890.
RA Aplin H.M., Hirst K.L., Crosby A.H., Dixon M.J.;
RT "Mapping of the human dentin matrix acidic phosphoprotein gene (DMP1)
RT to the dentinogenesis imperfecta type II critical region at chromosome
RT 4q21.";
RL Genomics 30:347-349(1995).
CC -!- FUNCTION: COULD BE INVOLVED IN THE INDUCTION OF MINERALIZATION OF
CC EXTRACELLULAR MATRIX.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ISOFORM 1 (SHOWN HERE) AND
CC ISOFORM 2; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN TOOTH PARTICULARLY IN
CC ODONTOBLAST, AMELOBLAST AND CEMENTOBLAST.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U89012; AAC51332.1; -;
CC EMBL; U34037; AAA97602.1; -;
CC EMBL; U65378; AAB87728.1; -;
CC MIM; 600980; -;
CC Extracellular matrix; Signal; Alternative splicing.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 513 DENTIN MATRIX ACIDIC PHOSPHOPROTEIN 1.
FT CARBOHYD 25 25 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 345 345 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 413 413 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 426 426 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 467 467 N-LINKED (GLCNAC...) (POTENTIAL).
FT SITE 364 366 CELL ATTACHMENT SITE (POTENTIAL).
FT VARSPIC 46 61 MISSING (IN ISOFORM 2).
FT CONFLICT 69 69 S -> C (IN REF. 2).
SQ SEQUENCE 513 AA; 55782 MW; 2C1FDE319A5D106F CRC64;

Query Match 74.4%; Score 29; DB 1; Length 513;
Best Local Similarity 62.5%; Pred. No. 49;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSTSSQSQ 8
Db 436 HSSSAESQ 443
|||||

```
CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN. BELONGS TO THE FES/FPS SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J02087; AAA43041.1; -.
CC PIR; A00651; TVMVGC.
CC HSP; P11362; IFGI.
CC INTERPRO; IPR000719; -.
CC INTERPRO; IPR000980; -.
CC INTERPRO; IPR001060; -.
CC INTERPRO; IPR001245; -.
CC PFAM; PF00611; FCH; 1.
CC PFAM; PF00017; SH2; 1.
CC PFAM; PF00069; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC PRINTS; PR00401; SH2DOMAIN.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS50001; SH2; 1.
CC Polyprotein; Transforming protein; Tyrosine-protein kinase; Oncogene;
CC Transferrase; ATP-binding; Phosphorylation; SH2 domain.
CC FT DOMAIN 247 336 SH2.
CC FT DOMAIN 348 609 PROTEIN KINASE.
CC FT NP_BIND 354 362 ATP (BY SIMILARITY).
CC FT BINDING 377 377 ATP (BY SIMILARITY).
CC FT ACT_SITE 470 470 BY SIMILARITY.
CC FT MOD_RES 500 500 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
CC SQ SEQUENCE 609 AA; 68769 MW; 53D4919757CF73A3 CRC64;

Query Match 76.9%; Score 30; DB 1; Length 609;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 HSTSSQSQ 8
    |||||
DB 194 HSTSSSEQ 201

RESULT 8
FES_FELCA STANDARD; PRT; 820 AA.
AC P14238;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE PROTO-ONCOGENE TYROSINE-PROTEIN KINASE FES/FPS (EC 2.7.1.112) (C-FES).
GN FES OR FPS.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 87198954.
RA Roebroek A.J.M.; Schalken J.A.; Onnekink C.; Bloemers H.P.J.;
RT van de Ven W.J.M.;
RT "Structure of the feline c-fes/fps proto-oncogene: genesis of a
RT retroviral oncogene.";
RL J. Virol. 61:2009-2016(1987).
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN. BELONGS TO THE FES/FPS SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M16705; AAA30808.1; JOINED.
CC EMBL; M16666; AAA30808.1; JOINED.
CC EMBL; M16667; AAA30808.1; JOINED.
CC EMBL; M16668; AAA30808.1; JOINED.
CC EMBL; M16669; AAA30808.1; JOINED.
CC EMBL; M16670; AAA30808.1; JOINED.
CC EMBL; M16671; AAA30808.1; JOINED.
CC EMBL; M16706; AAA30808.1; JOINED.
CC EMBL; M16672; AAA30808.1; JOINED.
CC EMBL; M16673; AAA30808.1; JOINED.
CC EMBL; M16674; AAA30808.1; JOINED.
CC EMBL; M16698; AAA30808.1; JOINED.
CC EMBL; M16700; AAA30808.1; JOINED.
CC EMBL; M16701; AAA30808.1; JOINED.
CC EMBL; M16702; AAA30808.1; JOINED.
CC EMBL; M16704; AAA30808.1; JOINED.
CC PIR; A27824; TVCTFF.
CC HSP; P11362; IFGI.
CC INTERPRO; IPR000719; -.
CC INTERPRO; IPR000980; -.
CC INTERPRO; IPR001060; -.
CC INTERPRO; IPR001245; -.
CC PFAM; PF00611; FCH; 1.
CC PFAM; PF00017; SH2; 1.
CC PFAM; PF00069; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS50001; SH2; 1.
CC Transferrase; Tyrosine-protein kinase; Proto-oncogene; ATP-binding;
CC Phosphorylation; SH2 domain.
CC FT DOMAIN 458 547 SH2.
CC FT DOMAIN 559 820 PROTEIN KINASE.
CC FT NP_BIND 565 573 ATP (BY SIMILARITY).
CC FT BINDING 588 588 ATP (BY SIMILARITY).
CC FT ACT_SITE 681 681 BY SIMILARITY.
CC FT MOD_RES 711 711 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
CC SQ SEQUENCE 820 AA; 92974 MW; F3A52B750236834E CRC64;

Query Match 76.9%; Score 30; DB 1; Length 820;
Best Local Similarity 75.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 HSTSSQSQ 8
    |||||
DB 405 HSTSSSEQ 412

RESULT 9
FES_HUMAN STANDARD; PRT; 822 AA.
AC P07332;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE PROTO-ONCOGENE TYROSINE-PROTEIN KINASE FES/FPS (EC 2.7.1.112) (C-FES).
GN FES OR FPS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
```

[3]
RN SEQUENCE OF 103-137 (MYELOMA PROTEIN NIG-65), AND CARBOHYDRATES.
RX MEDLINE; 82231401.
RA Takayasu T., Suzuki S., Kametani F., Takahashi N., Shinoda T.,
RT Okuyama T., Muneoka E.;
"Amino acid sequence of galactosamine-containing glycopeptides in the
RL hinge region of a human immunoglobulin D.";
RL Biochem. Biophys. Res. Commun. 105:1066-1071(1982).
[4]
RN SEQUENCE OF 158-383 (MYELOMA PROTEIN NIG-65).
RX MEDLINE; 8119406.
RA Shinoda T., Takahashi N., Takayasu T., Okuyama T., Shimizu A.;
RT "Complete amino acid sequence of the Fc region of a human delta
chain.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:785-789(1981).
[5]
RN CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE; 81133632.
RA Takayasu T., Takahashi N., Shinoda T.;
RT "Amino acid sequence and location of the three glycopeptides in the
RL Fc region of human immunoglobulin D.";
RL Biochem. Biophys. Res. Commun. 97:635-641(1980).
DR PIR; A02175; DHU.
DR MIM; 147170;
DR INTERPRO; IPR000495;
DR INTERPRO; IPR003006;
DR PFAM; PF00047; Ig; 3.
DR PROSITE; PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1 1
FT DISULFID 15 15 INTERCHAIN (WITH LIGHT CHAIN).
FT DISULFID 28 84 INTERCHAIN (WITH HEAVY CHAIN).
FT CARBOHYD 161 161
FT CARBOHYD 109 109
FT CARBOHYD 110 110
FT CARBOHYD 113 113
FT CARBOHYD 126 126
FT CARBOHYD 127 127
FT CARBOHYD 131 131
FT CARBOHYD 132 132
FT CARBOHYD 225 225
FT CARBOHYD 316 316
FT CARBOHYD 367 367
SQ SEQUENCE 383 AA; 42125 MW; DAIDE73519C76CID CRC64;

Query Match 76.9%; Score 30; DB 1; Length 383;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
Db 223 HNGSQSQ 230

RESULT 6
FES_FSVST STANDARD; PRT; 477 AA.
AC P00543;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE TYROSINE-PROTEIN KINASE TRANSFORMING PROTEIN FES (EC 2.7.1.112).
GN V-FES.
OS Feline sarcoma virus (strain Snyder-Theilen).
OC Viruses; Retroviridae; Retroviridae; Mammalian type C retroviruses.
RN [1]
RP SEQUENCE FROM N.A.
RA Hampe A., Laprevotte I., Galibert F., Fedele L.A., Sherr C.J.;
RX MEDLINE; 83050963.
RT "Nucleotide sequences of feline retroviral oncogenes (v-fes) provide
RL evidence for a family of tyrosine-specific protein kinase genes.";
Cell 30:775-785(1982).

CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-FES
CC POLYPROTEIN.
CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN. BELONGS TO THE FES/FPS SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J02088; CAB23266.1; ALT_SEQ.
DR PIR; A00652; TVMVCS.
DR HSSP; P11362; IFGI.
DR INTERPRO; IPR000719;
DR INTERPRO; IPR000980;
DR INTERPRO; IPR001245;
DR PFAM; PF00017; SH2; 1.
DR PFAM; PF00069; Pkinase; 1.
DR PRINTS; PRO0109; TYRKINASE.
DR PRINTS; PRO0401; SH2DOMAIN.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50001; SH2; 1.
KW Polypeptide; Transforming protein; Tyrosine-protein kinase; Oncogene;
KW Transferase; ATP-binding; Phosphorylation; SH2 domain.
FT DOMAIN 115 204
FT DOMAIN 216 477 PROTEIN KINASE.
FT NP_BIND 222 230 ATP (BY SIMILARITY).
FT BINDING 245 245 ATP (BY SIMILARITY).
FT ACT_SITE 338 338 BY SIMILARITY.
FT MOD_RES 368 368 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 477 AA; 53756 MW; BB87EDB47A3BE95 CRC64;

Query Match 76.9%; Score 30; DB 1; Length 477;
Best Local Similarity 75.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
Db 62 HSTSSEQ 69

RESULT 7
FES_FSVGA STANDARD; PRT; 609 AA.
ID FES_FSVGA
AC P00542;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE TYROSINE-PROTEIN KINASE TRANSFORMING PROTEIN FES (EC 2.7.1.112).
GN V-FES.
OS Feline sarcoma virus (strain Gardner-Arnstein) (Ga-FesV) (Gardner-
OS Arnstein feline leukemia oncovirus B).
OC Viruses; Retroviridae; Retroviridae; Mammalian type C retroviruses.
RN [1]
RP SEQUENCE FROM N.A.
RA Hampe A., Laprevotte I., Galibert F., Fedele L.A., Sherr C.J.;
RT "Nucleotide sequences of feline retroviral oncogenes (v-fes) provide
RL evidence for a family of tyrosine-specific protein kinase genes.";
Cell 30:775-785(1982).
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-FES
CC POLYPROTEIN.

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39481; AAA85862.1; -.
DR EMBL; U43491; AAC49488.1; -.
DR EMBL; Z74916; CAA99196.1; -.
DR HSSP; P05067; IBA6.
DR SGD; S0005534; SLG1.
DR INTERPRO; IPR002889; -.
DR PFAM; PF01822; WSC; 1.
KW Glycoprotein; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 378 SLG1 PROTEIN.
FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 378 AA; 39270 MW; EEE164F2374CCCE3 CRC64;

Query Match 79.5%; Score 31; DB 1; Length 378;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 STSSQSQ 8
Db 200 STSSQSQ 206
|||||

RESULT 4
SLA2_YEAST STANDARD; PRT; 968 AA.
AC P33338; Q02434;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE SLA2 PROTEIN (TRANSMEMBRANE PROTEIN MOP2).
GN SLA2 OR END4 OR MOP2 OR UFG1 OR YNL243W OR N1102.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomycetes.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DDY 228;
RX MEDLINE; 93328765.
RA Holtzman D.A., Yang S., Drubin D.G.;
RT "Synthetic-lethal interactions identify two novel genes, SLA1 and
RT SLA2, that control membrane cytoskeleton assembly in Saccharomycetes
RT cerevisiae.";
RL J. Cell Biol. 122:635-644(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95204480.
RA Na S., Hincapié M., McCusker J.H., Haber J.E.;
RT "MOP2 (SLA2) affects the abundance of the plasma membrane H(+)-ATPase
RT of Saccharomycetes cerevisiae.";
RL J. Biol. Chem. 270:6815-6823(1995).
RN [3]
RP SEQUENCE OF 327-968 FROM N.A.
RC STRAIN-117-8A;
RA Yoon H., Donahue T.F.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Saiz J.E., Baladron V., del Rey F.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 819-968 FROM N.A.
RX MEDLINE; 97051596.
RA Pandolfo D., de Antoni A., Lanfranchi G., Valle G.;
RT "The DNA sequence of cosmid 14-5 from chromosome XIV reveals 21 open
RT reading frames including a novel gene encoding a globin-like
RT domain.";
RL Yeast 12:1071-1076(1996).
CC -1- FUNCTION: REQUIRED FOR CELLULAR MORPHOGENESIS AND POLARIZATION OF
CC THE CORTICAL CYTOSKELETON. IT MIGHT ACT IN CONCERT WITH PROTEINS
```

```
CC SUCH AS CDC42 AND CDC43 TO LIMIT THE REGION OF CORTICAL PATCH
CC FORMATION TO THE CORTEX OF THE BUD. REQUIRED FOR THE ACCUMULATION
CC AND/OR MAINTENANCE OF PLASMA MEMBRANE H(+)-ATPASE ON THE CELL
CC SURFACE.
CC -1- SIMILARITY: TO C.ELEGANS ZK370.3 AND TO THE C-TERMINUS OF TALIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z22811; CAA80464.1; -.
DR EMBL; L12352; AAA74726.1; -.
DR EMBL; U07938; AAA19161.1; -.
DR EMBL; Z71519; CAA96149.1; -.
DR EMBL; Z71518; CAA96148.1; -.
DR EMBL; Z69381; CAA93355.1; -.
DR PIR; S36354; S36354.
DR PIR; B40673; B40673.
DR SGD; S0005187; SLA2.
DR INTERPRO; IPR001026; -.
DR INTERPRO; IPR002558; -.
DR PFAM; PF01417; ENTH; 1.
DR PFAM; PF01608; L_LWEO; 1.
KW Cytoskeleton; Actin-binding; Transmembrane.
FT TRANSMEM 772 791 POTENTIAL.
FT CONFLICT 52 52 P -> A (IN REF. 2).
FT CONFLICT 344 344 A -> R (IN REF. 4).
FT CONFLICT 560 560 A -> R (IN REF. 3).
FT CONFLICT 887 968 EHCKSDVTDACKRSLGCHVGMIEDDHSTSQQQPLDFTTSEH
FT TLKTAEMEQVEILKLEQSLNARKRLGEIRRHAYNQDDDD
FT -> NTPVKMLPMHAEVWVTTSNV (IN REF. 3).
SQ SEQUENCE 968 AA; 108910 MW; E592E09D8040C0E9 CRC64;

Query Match 79.5%; Score 31; DB 1; Length 968;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HSTSSQSQ 8
Db 912 HSTSSQSQ 919
|||||

RESULT 5
DTC_HUMAN STANDARD; PRT; 383 AA.
ID P01880;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG DELTA CHAIN C REGION.
GN IGHD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Plrimates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE OF 1-162 (MYELOMA PROTEIN WAH).
RX MEDLINE; 82082419.
RA Putnam F.W., Takahashi N., Tetaert D., Debuire B., Lin L.-C.;
RT "Amino acid sequence of the first constant region domain and the
RT hinge region of the delta heavy chain of human IgD.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:6168-6172(1981).
RN [2]
RP SEQUENCE OF 158-383 (MYELOMA PROTEIN WAH).
RX MEDLINE; 81223768.
RA Lin L.-C., Putnam F.W.;
RT "Primary structure of the Fc region of human immunoglobulin D;
RT implications for evolutionary origin and biological function.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:504-508(1981).
```


FT TRANSMEM 274 295 4 (POTENTIAL).
 FT DOMAIN 296 313 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 314 334 5 (POTENTIAL).
 FT DOMAIN 335 361 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 362 380 6 (POTENTIAL).
 FT DOMAIN 381 391 7 (POTENTIAL).
 FT TRANSMEM 392 414 7 (POTENTIAL).
 FT DOMAIN 415 546 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 546 AA; 61800 MW; 2825AE4040313527 CRC64;

Query Match 100.0%; Score 39; DB 1; Length 546;
 Best Local Similarity 100.0%; Pred. No. 0.34;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSTSSQSQ 8
 Db 454 HSTSSQSQ 461
 |||||

RESULT 2
 PTR2_HUMAN STANDARD; PRT; 550 AA.
 ID PTR2_HUMAN STANDARD; PRT; 550 AA.
 AC P49190;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PARATHYROID HORMONE RECEPTOR PRECURSOR (PTH2 RECEPTOR).
 GN PTHR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE-BRAIN;
 RX MEDLINE; 95318121.
 RA Usdin T.B., Gruber C., Bonner T.I.;
 RT "Identification and functional expression of a receptor selectively
 RT recognizing parathyroid hormone, the PTH2 receptor.";
 RL J. Biol. Chem. 270:15455-15458(1995).
 RN [2]
 RP SEQUENCE OF 26-40 AND 306-550 FROM N.A.
 RX MEDLINE; 97079671.
 RA Usdin T.B., Modi W., Bonner T.I.;
 RT "Assignment of the human PTH2 receptor gene (PTH2) to chromosome 2q33
 RT by fluorescence in situ hybridization.";
 RL Genomics 37:140-141(1996).
 CC -!- FUNCTION: THIS IS A SPECIFIC RECEPTOR FOR PARATHYROID HORMONE. THE
 CC ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
 CC ADENYLYL CYCLASE.
 CC -!- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN BRAIN AND PANCREAS.
 CC ALSO EXPRESSED IN THE TESTIS.
 CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U25128; AAC50157.1;
 CC EMBL; U47124; AAA96796.1;
 CC EMBL; U47129; AAC50767.1;
 CC EMBL; U47125; AAC50767.1; JOINED.
 CC EMBL; U47126; AAC50767.1; JOINED.
 CC EMBL; U47127; AAC50767.1; JOINED.
 CC EMBL; U47128; AAC50767.1; JOINED.

DR GCRDB; GCR_2003;
 DR MM; 601469;
 DR INTERPRO; IPRO00832;
 DR PFAM; PF00002; 7tm2.1;
 DR PRINTS; PR00249; GPCRSECRETIN
 DR PROSITE; PS00649; G-PROTEIN RECEPTOR_F2_1; 1.
 DR PROSITE; PS00650; G-PROTEIN RECEPTOR_F2_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 550
 FT DOMAIN 27 145 PARATHYROID HORMONE RECEPTOR.
 FT TRANSMEM 146 169 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 170 176 1 (POTENTIAL).
 FT TRANSMEM 177 196 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 197 237 2 (POTENTIAL).
 FT TRANSMEM 238 260 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 261 275 3 (POTENTIAL).
 FT TRANSMEM 276 297 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 298 316 4 (POTENTIAL).
 FT TRANSMEM 317 337 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 338 364 5 (POTENTIAL).
 FT TRANSMEM 365 383 6 (POTENTIAL).
 FT DOMAIN 384 394 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 395 417 7 (POTENTIAL).
 FT DOMAIN 418 550 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 550 AA; 62235 MW; 2ADD14DBA68A9BF8 CRC64;

Query Match 100.0%; Score 39; DB 1; Length 550;
 Best Local Similarity 100.0%; Pred. No. 0.35;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSTSSQSQ 8
 Db 457 HSTSSQSQ 464
 |||||

RESULT 3
 SLGL_YEAST
 ID SLGL_YEAST STANDARD; PRT; 378 AA.
 AC P54867;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE SLGL PROTEIN PRECURSOR.
 GN SLGL OR YOR008C OR UNF378.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomycetes.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA de Beggan G., Bergey-Aullo P., Barthe C., Louvet O.,
 RA Peypouquet M.F., Morel C., Doignon F., Crouzet M.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 97051599.
 RA Sterky F., Holmberg A., Pettersson B., Uhlen M.;
 RT "The sequence of a 30 kb fragment on the left arm of chromosome XV
 RT from Saccharomyces cerevisiae reveals 15 open reading frames, five of
 RT which correspond to previously identified genes.";
 RL Yeast 12:1091-1095(1996).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 09:03:53 ; Search time 58.45 Seconds
(without alignments)
4.373 Million cell updates/sec

Title: us-09-236-468a-2_copy_457_464

Perfect score: 39
Sequence: 1 HSTSSQSQ 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 87993

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	100.0	546	1 PTH2_RAT	P70555 rattus norv
2	39	100.0	550	1 PTH2_HUMAN	P49190 homo sapien
3	31	79.5	378	1 SLG1_YEAST	P54867 saccharomyc
4	31	79.5	968	1 SLA2_YEAST	P33338 saccharomyc
5	30	76.9	383	1 DTC_HUMAN	P01880 homo sapien
6	30	76.9	477	1 FES_FSVST	P00543 feline sarc
7	30	76.9	609	1 FES_FSVGA	P14238 felis silve
8	30	76.9	820	1 FES_FELCA	P07332 homo sapien
9	30	76.9	822	1 FES_HUMAN	Q13316 homo sapien
10	29	74.4	513	1 DMPL_HUMAN	Q09729 schizosacch
11	29	74.4	596	1 YAMC_SCHPO	P46463 pichia past
12	29	74.4	1157	1 PEX1_PICPA	O91690 xenopus lae
13	29	74.4	1360	1 GLI1_XENLA	O08565 rattus norv
14	28	71.8	349	1 CCR4_RAT	P20930 homo sapien
15	28	71.8	416	1 FTLA_HUMAN	P20417 rattus norv
16	28	71.8	432	1 PTNL_RAT	Q15516 homo sapien
17	28	71.8	846	1 CLOC_HUMAN	P53104 saccharomyc
18	28	71.8	897	1 APG1_YEAST	P53935 saccharomyc
19	28	71.8	1240	1 YNJ1_YEAST	P38243 saccharomyc
20	27	69.2	211	1 YBS1_YEAST	P44697 haemophilus
21	27	69.2	269	1 THID_HAEIN	P27021 human respi
22	27	69.2	297	1 VGLG_HRSV2	P27022 human respi
23	27	69.2	297	1 VGLG_HRSV3	P27024 human respi
24	27	69.2	298	1 VGLG_HRSV5	P30423 human respi
25	27	69.2	298	1 VGLG_HRSVA	P20895 human respi
26	27	69.2	298	1 VGLG_HRSVL	P11113 physarum po
27	27	69.2	332	1 SRA_PHYPO	P16864 xenopus lae
28	27	69.2	334	1 CXA2_XENLA	O00476 homo sapien
29	27	69.2	401	1 NPT4_HUMAN	P37137 xenopus lae
30	27	69.2	402	1 LHX5_XENLA	O91178 oryzias lat
31	27	69.2	428	1 GPRX_ORYLA	P47980 drosophila
32	27	69.2	437	1 TIS1_DROME	P23630 bacillus su
33	27	69.2	441	1 DCDA_BAGSU	

ALIGNMENTS

RESULT	1				
PTH2_RAT					
ID	PTH2_RAT	STANDARD;	PRT;	546 AA.	
AC	P70555;				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	30-MAY-2000 (Rel. 39, Last annotation update)				
DE	PARATHYROID HORMONE RECEPTOR PRECURSOR (PTH2 RECEPTOR).				
GN	PTH2.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 96426194.				
PA	Usdin T.B., Bonner T.I., Harta G., Mezey E.;				
RT	"Distribution of parathyroid hormone-2 receptor messenger ribonucleic acid in rat.";				
RT	Endocrinology 137:4285-4297(1996).				
CC	-!- FUNCTION: THIS IS A SPECIFIC RECEPTOR FOR PARATHYROID HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYL CYCLASE. PTH2 MAY BE RESPONSIBLE FOR PTH EFFECTS IN A NUMBER OF PHYSIOLOGICAL SYSTEMS. IT MAY PLAY A SIGNIFICANT ROLE IN PANCREATIC FUNCTION. PTH2 PRESENCE IN NEURONS INDICATES THAT IT MAY FUNCTION AS A NEUROTRANSMITTER RECEPTOR.				
CC	-!- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN BRAIN, ARTERIAL AND CARDIAC ENDOTHELIUM. FOUND AS WELL IN SPERM. IN THE HEAD OF THE EPIDIDYMIS. LOWER EXPRESSION IS FOUND IN VASCULAR SMOOTH MUSCLE, EXOCRINE PANCREAS, TESTIS AND PLACENTA.				
CC	-!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	EMBL; U55836; AAC52849.1; -.				
DR	CCRD; GCR_1413; -.				
DR	INTERPRO: IPR000832; -.				
DR	PFAM: PF00002; 7tm.2; 1.				
DR	PRINIS; PR00249; GPCRSECRETIN.				
DR	PROSITE; PS00649; G_PROTEIN_RECEP_F2.1; 1.				
DR	PROSITE; PS00650; G_PROTEIN_RECEP_F2.2; FALSE_NEG.				
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.				
FT	SIGNAL 1 24 POTENTIAL.				
FT	CHAIN 25 546 PARATHYROID HORMONE RECEPTOR.				
FT	DOMAIN 27 143 EXTRACELLULAR (POTENTIAL).				
FT	TRANSMEM 144 167 1 (POTENTIAL).				
FT	DOMAIN 168 174 CYTOPLASMIC (POTENTIAL).				
FT	TRANSMEM 175 194 2 (POTENTIAL).				
FT	DOMAIN 195 235 EXTRACELLULAR (POTENTIAL).				
FT	TRANSMEM 236 258 3 (POTENTIAL).				
FT	DOMAIN 259 273 CYTOPLASMIC (POTENTIAL).				

O35618 mus musculu
P48565 saccharomyc
P42826 saccharomyc
P38700 saccharomyc
P03072 polyomaviru
P14999 polyomaviru
P03071 polyomaviru
P32891 saccharomyc
P22966 homo sapien
P34588 caenorhabdi
P28840 rattus norv
P29120 homo sapien

THIS PAGE BLANK (USPTO)

Query Match 79.5%; Score 31; DB 2; Length 2186;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
||:||||
Db 2152 HSSSRQ 2159

RESULT 15
A35938
profilaggrin - human (fragments)
C:Species: Homo sapiens (man)
C:Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C:Accession: A35938
R:Gan, S.Q.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.
Biochemistry 29, 9432-9440, 1990
A:Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A:Reference number: A35938; MUID:91064347
A:Accession: A35938
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-2248 <GAN>
A:Cross-references: GB:J02929
C:Genetics:
A:Gene: GDB:FLG
A:Cross-references: GDB:I19912; OMIM:135940
A:Map position: lq21-lq21
C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C:Keywords: EF hand; epidermis; polymorphism; tandem repeat
F:246-569/Region: filaggrin repeat
F:570-893/Region: filaggrin repeat
F:1074-1397/Region: filaggrin repeat
F:1573-1896/Region: filaggrin repeat

Query Match 79.5%; Score 31; DB 2; Length 2248;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
|||||
Db 284 HSTSQQGQ 291

Search completed: November 8, 2000, 08:53:34
Job time: 365 sec

Query Match 79.5%; Score 31; DB 2; Length 786;
Best Local Similarity 75.0%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
|||||
Db 275 HNTSSDSQ 282

RESULT 12
S63211
cytoskeleton assembly control protein SLA2 - yeast (Saccharomyces cerevisiae)
N:Alternate names: MOP2 protein; protein N1102; protein YNL243w; UFG1 protein
C:Species: Saccharomyces cerevisiae
C>Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 29-Oct-1999
C:Accession: S63211; S63209; B40673; S61932; S64652; S67355; S72073; S36354; S42278
R:Salz, J.E.; Baladron, V.; Del Rey, F.
Submitted to the Protein Sequence Database, April 1996
A:Reference number: S63210
A:Accession: S63211
A:Molecule type: DNA
A:Residues: 1-968 <SAI>
A:Cross-references: EMBL:Z71519; NID:gl302279; PIDN:CAA96149.1; PID:e239710; PID:gl30228
R:Pandolfo, D.; De Antoni, A.; Lanfranchi, G.; Valle, G.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63188
A:Accession: S63209
A:Molecule type: DNA
A:Residues: 820-968 <PAM>
A:Cross-references: EMBL:Z71519
R:Holtzman, D.A.; Yang, S.; Drubin, D.G.
J. Cell Biol. 122, 635-644, 1993
A:Title: Synthetic-lethal interactions identify two novel genes, SLA1 and SLA2, that control cell growth
A:Reference number: A40673; MUID:93328765
A:Accession: B40673
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-343, 'A', 'A', '345-968 <HOL>
A:Cross-references: GB:222811; NID:g311412; PIDN:CAA80464.1; PID:g311413
R:Yoon, H.; Donahue, T.F.
submitted to the EMBL Data Library, March 1994
A:Description: DNA sequence of ufg1 gene in yeast.
A:Reference number: S61932
A:Accession: S61932
A:Molecule type: DNA
A:Residues: 327-343, 'A', '345-559, 'R', '561-886, 'NTVP', '891, 'MLPMHAEVWVTTSW', '904 <YOO>
A:Cross-references: EMBL:U07938; NID:g501047; PIDN:AAA19161.1; PID:g501048
R:Na, S.; Hincapié, M.; McCusker, J.H.; Haber, J.E.
J. Biol. Chem. 270, 6815-6823, 1995
A:Title: MOP2 (SLA2) affects the abundance of the plasma membrane H(+) -ATPase of Saccharomyces cerevisiae
A:Reference number: S64652; MUID:95204480
A:Accession: S64652
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-51, 'A', '53-343, 'A', '345-968 <NAS>
A:Cross-references: EMBL:L12352; NID:g870731; PIDN:AAA74726.1; PID:g870732
R:Pandolfo, D.; De Antoni, A.; Lanfranchi, G.; Valle, G.
submitted to the EMBL Data Library, February 1996
A:Description: DNA sequence of cosmid 14-5 from chromosome XIV.
A:Reference number: S67355
A:Accession: S67355
A:Molecule type: DNA
A:Residues: 820-968 <PAM>
A:Cross-references: EMBL:Z69381; NID:gl183970; PIDN:CAA93355.1; PID:e221951; PID:gl18397
R:Pandolfo, D.; De Antoni, A.; Lanfranchi, G.; Valle, G.
Yeast 12, 1071-1076, 1996
A:Title: The DNA sequence of cosmid 14-5 from chromosome XIV reveals 21 open reading frames
A:Reference number: S72073; MUID:97051596

A:Accession: S72073
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 820-968 <PAF>
A:Cross-references: EMBL:Z69381; NID:gl183970; PIDN:CAA93355.1; PID:e221951; PID:gl18397
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996
C:Genetics:
A:Gene: SGD:SLA2; UFG1; END4; MOP2
A:Cross-references: MIPS:YNL243w; SGD:S0005187
A:Map position: 14L
C:Function:
A:Description: essential for the assembly and function of the cortical cytoskeleton
C:Keywords: transmembrane protein
F:771-787/Domain: transmembrane #status predicted <TMM>

Query Match 79.5%; Score 31; DB 2; Length 968;
Best Local Similarity 75.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
|||||
Db 912 HSTSQQQQ 919

RESULT 13
T18543
probable cell-adhesion protein MFB1 - shiitake mushroom
C:Species: Lentinula edodes (shiitake mushroom)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18543
R:Kondoh, O.; Muto, A.; Kajiwarra, S.; Takagi, J.; Saito, Y.; Shishido, K.
Gene 154, 31-37, 1995
A:Title: A fruiting body-specific cDNA, mfbac, from the mushroom Lentinus edodes encodes a cell-adhesion protein
A:Reference number: Z18360; MUID:95172398
A:Accession: T18543
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2140 <KON>
A:Cross-references: EMBL:D14487; NID:d1095773; PID:d1019638; PIDN:BAA18903.1
A:Experimental source: strain FMC2
C:Genetics:
A:Gene: MFB1

Query Match 79.5%; Score 31; DB 2; Length 2140;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
|||||
Db 743 HSTASSSQ 750

RESULT 14
T13169
tigrin - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 13-Aug-1999
C:Accession: T13169
R:Fogerty, F.J.; Fessler, L.I.; Bunch, T.A.; Yaron, Y.; Parker, C.G.; Nelson, R.E.; B
Development 120, 1747-1758, 1994
A:Title: Tigrin, a novel Drosophila extracellular matrix protein that functions as a tigrin
A:Reference number: Z17625; MUID:95009506
A:Accession: T13169
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2186 <FOG>
A:Cross-references: EMBL:U09506; NID:g493069; PID:g493070; PIDN:AAA56958.1
C:Function:
A:Description: functions as a ligand for Drosophila alpha PS2 beta PS integrins

Qy 2 STSSQSQ 8
|||||
Db 200 STSSQSQ 206

RESULT 7

C42094
bw3 protein - smut fungus (Ustilago maydis) (fragment)
C:Species: Ustilago maydis (corn smut)
C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 24-Sep-1999
C:Accession: C42094
R:Gillissen, B.; Bergemann, J.; Sandmann, C.; Schroeer, B.; Boelker, M.; Kahmann, R.
Cell 68, 647-657, 1992
A:Title: A two-component regulatory system for self/non-self recognition in Ustilago maydis
A:Reference number: A42094; MUID:92154679
A:Accession: C42094
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-405 <GIL>
A:Cross-references: GB:M84180; NID:g170576; PIDN:AAA34222.1; PID:g170577
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:137-193/Domain: homeobox homology <HOX>

Query Match 79.5%; Score 31; DB 2; Length 405;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
|||||
Db 292 HSTSSSSE 299

RESULT 8

T19682
hypothetical protein C33D9.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19682
R:Lloyd, C.

submitted to the EMBL Data Library, December 1995

A:Reference number: Z19162

A:Accession: T19682

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-604 <WIL>

A:Cross-references: EMBL:Z68159; PIDN:CAA92286.1; GSPDB:GN00022; CESP:C33D9.8

A:Experimental source: clone C33D9

C:Genetics:

A:Gene: CESP:C33D9.8

A:Map position: 4

A:Introns: 215/2; 516/1

Query Match 79.5%; Score 31; DB 2; Length 604;
Best Local Similarity 85.7%; Pred. No. 43;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HSTSSQS 7
|||||
Db 7 HSTSES 13

RESULT 9

T26232
hypothetical protein W06D12.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26232
R: Basham, V.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z20177

A:Accession: T26232

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-631 <WIL>

A:Cross-references: EMBL:Z82073; PIDN:CAB04926.1; GSPDB:GN00023; CESP:W06D12.5

A:Experimental source: clone W06D12

C:Genetics:

A:Gene: CESP:W06D12.5

A:Map position: 5

A:Introns: 4/3; 50/1; 89/2; 153/1; 244/3; 285/3; 333/1; 384/2; 509/1; 553/3

Query Match 79.5%; Score 31; DB 2; Length 631;
Best Local Similarity 75.0%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
|||||
Db 613 HRTASQSQ 620

RESULT 10

T50302
stress activated map kinase interacting protein [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000
C:Accession: T50302
R:McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.; Saunders, D.; Harris, D.
submitted to the EMBL Data Library, January 2000

A:Reference number: Z25059

A:Accession: T50302

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-665 <MCD>

A:Cross-references: EMBL:AL136521; PIDN:CAB66311.1; GSPDB:GN00066; SPDB:SPAPYUG7.02C

A:Experimental source: strain 972h(-); clone plasmid pYUG7

C:Genetics:

A:Gene: SPDB:SPAPYUG7.02C

A:Map position: 1

A:Introns: 8/2; 636/2

Query Match 79.5%; Score 31; DB 2; Length 665;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 STSSQSQ 8
|||||
Db 419 STSSQSQ 425

RESULT 11

T39585

ubiquitin protein ligase - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000

C:Accession: T39585

R:Voickaert, G.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z21865

A:Accession: T39585

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-786 <VOL>

A:Cross-references: EMBL:Z99759; PIDN:CAB16903.1; GSPDB:GN00067; SPDB:SPBC16E9.11C

A:Experimental source: strain 972h-; cosmid c16E9

C:Genetics:

A:Gene: SPDB:SPBC16E9.11C

A:Map position: 2

A:Introns: 60/2

C:Superfamily: yeast ubiquitin--protein ligase; WW repeat homology

C;Superfamily: collagenase prtc

Query Match 82.1%; Score 32; DB 2; Length 451;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

QY 1 HSTSSQSQ 8
|||: |||
Db 364 HSTAKSQSQ 371

RESULT 3

B81056
proteinase, probable NMB1664 [imported] - Neisseria meningitidis (group B strain MD58)
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 26-May-2000
C;Accession: B81056
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizza, M.;
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755
A;Accession: B81056
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-451 <TET>
A;Cross-references: GB:AE002517; GB:AE002098; NID:g7226912; PIDN:AAF42013.1; PID:g722691
A;Experimental source: serogroup B, strain NMD58
C;Genetics:
A;Gene: NMB1664
C;Superfamily: collagenase prtc

Query Match 82.1%; Score 32; DB 2; Length 451;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HSTSSQSQ 8
|||: |||
Db 364 HSTAKSQSQ 371

RESULT 4

T13716
bazooka gene protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 13-Aug-1999
C;Accession: T13716
R;Kuchinke, U.; Grawe, F.; Knust, E.
submitted to the EMBL Data Library, November 1998
A;Description: Control of spindle orientation in Drosophila by the Par-3-related PDZ-dom
A;Reference number: Z117708
A;Accession: T13716
A;Status: preliminary; translated from GR/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1464 <KUC>
A;Cross-references: EMBL:AJ130871; NID:el363519; PID:el363520; PIDN:CAA10224.1
C;Genetics:
A;Gene: bazooka
A;Map position: X

Query Match 82.1%; Score 32; DB 2; Length 1464;
Best Local Similarity 87.5%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HSTSSQSQ 8
|||: |||
Db 1320 HSTSSGSGQ 1327

RESULT 5

T02905
hypothetical protein T13J8.150 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 24-Mar-1999
C;Accession: T02905
R;Bevan, M.; Pohl, T.; Weizenegger, T.; Hohelsel, J.; Mewes, H.W.; Mayer, K.F.X.; Sch
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z14766
A;Accession: T02905
A;Molecule type: DNA
A;Residues: 1-359 <BEV>
A;Cross-references: EMBL:AL035524
A;Experimental source: cultivar Columbia; BAC clone T13J8
C;Genetics:
A;Map position: 4
A;Introns: 60/2; 122/2; 295/2
A;Note: T13J8.150

Query Match 79.5%; Score 31; DB 2; Length 359;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSTSSQSQ 7
|||: |||
Db 330 HNTSSQSQ 336

RESULT 6

S61992
SLG1 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein O2545; protein UNF378; protein YOR008c
C;Species: Saccharomyces cerevisiae
C;Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 17-Mar-1999
C;Accession: S61992; S66873; S72141
R;Sterky, F.; Uhlen, M.
submitted to the EMBL Data Library, December 1995
A;Reference number: S61981
A;Accession: S61992
A;Molecule type: DNA
A;Residues: 1-378 <STE>
A;Cross-references: EMBL:U43491; NID:g1150992; PID:g1151004
R;Pettersson, B.; Sterky, F.; Uhlen, M.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66882
A;Accession: S66873
A;Molecule type: DNA
A;Residues: 1-378 <PET>
A;Cross-references: EMBL:Z74916; NID:g1420101; PID:e252318; PID:g1420102; MIPS:YOR008
R;Sterky, F.; Holmberg, A.; Pettersson, B.; Uhlen, M.
Yeast 12, 1091-1095, 1996
A;Title: The sequence of a 30 kb fragment on the left arm of chromosome XV from Sacch
A;Reference number: S72130; MUID:97051599
A;Accession: S72141
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-378 <STW>
A;Cross-references: EMBL:U43491; NID:g1150992; PID:g1151004
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995
C;Genetics:
A;Gene: SGD:SLG1
A;Cross-references: SGD:S0005534; MIPS:YOR008c
A;Map position: 15R
C;Keywords: transmembrane protein
F;270-286/Domain: transmembrane #status predicted <TM>

Query Match 79.5%; Score 31; DB 2; Length 378;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:53:32 ; Search time 99.87 Seconds
(without alignments)
5.083 Million cell updates/sec

Title: US-09-236-468A-2_COPY_457_464

Perfect score: 39

Sequence: 1 HSTSOSQ 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 182106

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- PIR_65:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	39	100.0	550	2 A57519	parathyroid hormon
2	32	82.1	451	2 C81820	probable proteinase
3	32	82.1	451	2 B81056	proteinase, probab
4	32	82.1	1464	2 T13716	bazooka gene prote
5	31	79.5	359	2 T02905	hypothetical prote
6	31	79.5	378	2 S61992	SLG1 protein - yea
7	31	79.5	405	2 C42094	bw3 protein - smut
8	31	79.5	604	2 T19682	hypothetical prote
9	31	79.5	631	2 T26232	hypothetical prote
10	31	79.5	665	2 T50302	stress activated -
11	31	79.5	786	2 T39585	ubiquitin proteli:
12	31	79.5	968	2 S63211	cytoskeleton assem
13	31	79.5	2140	2 T18543	probable cell-adhe
14	31	79.5	2186	2 T13169	tiggrin - fruit fl
15	31	79.5	2248	2 A35938	profilaggrin - hum
16	30	76.9	268	2 B30819	interferon-regulat
17	30	76.9	383	1 DHHU	Ig delta chain C r
18	30	76.9	477	1 TVMVCS	protein-tyrosine k
19	30	76.9	509	2 T17597	Ig delta chain (WI
20	30	76.9	571	2 T06737	hypothetical prote
21	30	76.9	609	1 TVMVGC	protein-tyrosine k
22	30	76.9	802	2 T23295	hypothetical prote
23	30	76.9	820	1 TVCTFF	protein-tyrosine k
24	30	76.9	822	1 TVHUFF	protein-tyrosine k
25	30	76.9	860	2 T23296	hypothetical prote
26	30	76.9	1469	2 T19168	hypothetical prote
27	29	74.4	220	2 S65477	ras-homolog Gprase
28	29	74.4	520	2 S74497	hypothetical prote
29	29	74.4	575	2 T01552	hypothetical prote

ALIGNMENTS

RESULT 1

A57519 parathyroid hormone receptor 2 precursor - human

N;Alternate names: PTH2 receptor

C;Species: Homo sapiens (man)

C;Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 24-Nov-1999

C;Accession: A57519

R;Usdin, T.B.; Gruber, C.; Bonner, T.I.

J. Biol. Chem. 270, 15455-15458, 1995

A;Title: Identification and functional expression of a receptor selectively recognizi

A;Reference number: A57519; MUID:95318121

A;Accession: A57519

A;Status: preliminary; nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-550 <USD>

A;Cross-references: GB:U25128; NID:g887966; PIDN:AAC50157.1; PID:g887967

C;Genetics:

A;Gene: GDB:PTHR2; PTHR2R

A;Cross-references: GDB:731977; OMIM:601469

A;Map position: 2q33-2q33

C;Superfamily: glucagon receptor

C;Keywords: hormone receptor

Query Match 100.0%; Score 39; DB 2; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSTSOSQ 8

Db 457 HSTSOSQ 464

RESULT 2

C81820 probable proteinase NMA1923 [imported] - Neisseria meningitidis (group A strain Z2491

C;Species: Neisseria meningitidis

C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 19-May-2000

C;Accession: C81820

R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000

A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491

A;Reference number: A81775; MUID:20222556

A;Accession: C81820

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-451 <PAR>

A;Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85144.1; PID:g738

A;Experimental source: serogroup A, strain Z2491

C;Genetics:

A;Gene: NMA1923

THIS PAGE BLANK (USPTO)

Search completed: November 8, 2000, 08:49:22
Job time: 115 sec

Query Match 69.2%; Score 27; DB 1; Length 732;
 Best Local Similarity 71.4%; Pred. No. 5.7e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HSTSSQ 7
 | | | | |
 Db 39 HGTSSQA 45

RESULT 13

US-08-712-241-2
 ; Sequence 2, Application US/08712241
 ; Patent No. 5789564
 ; GENERAL INFORMATION:
 ; APPLICANT: SEIDAH, NABIL G.
 ; TITLE OF INVENTION: DEVELOPMENT OF RESEARCH,
 ; TITLE OF INVENTION: DIAGNOSTIC AND PRODUCTION TOOLS FOR PRO-HORMONE
 ; TITLE OF INVENTION: CONVERTASES
 ; NUMBER OF SEQUENCES: 28
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: QUARLES & BRADY
 ; STREET: 411 EAST WISCONSIN AVENUE
 ; CITY: MILWAUKEE
 ; STATE: WISCONSIN
 ; COUNTRY: U.S.A.
 ; ZIP: 53202-4497
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 720 kb diskette
 ; COMPUTER: IBM PS/2, Model 30
 ; OPERATING SYSTEM: PC-DOS 3.30
 ; SOFTWARE: WordPerfect 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/712,241
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/529,785
 ; FILING DATE: OCTOBER 20, 1992
 ; APPLICATION NUMBER: US/07/963,535A
 ; FILING DATE: OCTOBER 20, 1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: JEAN C. BAKER
 ; REGISTRATION NUMBER: 35,433
 ; REFERENCE/DOCKET NUMBER: 20-702-9001-7
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (414) 277-5000
 ; TELEFAX: (414) 277-5591
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 753 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-712-241-2

Query Match 69.2%; Score 27; DB 1; Length 753;
 Best Local Similarity 62.5%; Pred. No. 5.9e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HSTSSQ 8
 | | | | |
 Db 592 HGTSSQPE 599

RESULT 14

US-08-712-241-3
 ; Sequence 3, Application US/08712241
 ; Patent No. 5789564

; GENERAL INFORMATION:
 ; APPLICANT: SEIDAH, NABIL G.
 ; APPLICANT: CHR TIEN, MICHEL
 ; TITLE OF INVENTION: DEVELOPMENT OF RESEARCH,
 ; TITLE OF INVENTION: DIAGNOSTIC AND PRODUCTION TOOLS FOR PRO-HORMONE
 ; TITLE OF INVENTION: CONVERTASES
 ; NUMBER OF SEQUENCES: 28
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: QUARLES & BRADY
 ; STREET: 411 EAST WISCONSIN AVENUE
 ; CITY: MILWAUKEE
 ; STATE: WISCONSIN
 ; COUNTRY: U.S.A.
 ; ZIP: 53202-4497
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 720 kb diskette
 ; COMPUTER: IBM PS/2, Model 30
 ; OPERATING SYSTEM: PC-DOS 3.30
 ; SOFTWARE: WordPerfect 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/712,241
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/529,785
 ; FILING DATE: OCTOBER 20, 1992
 ; APPLICATION NUMBER: US/07/963,535A
 ; FILING DATE: OCTOBER 20, 1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: JEAN C. BAKER
 ; REGISTRATION NUMBER: 35,433
 ; REFERENCE/DOCKET NUMBER: 20-702-9001-7
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (414) 277-5000
 ; TELEFAX: (414) 277-5591
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 753 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-712-241-3

Query Match 69.2%; Score 27; DB 1; Length 753;
 Best Local Similarity 62.5%; Pred. No. 5.9e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HSTSSQ 8
 | | | | |
 Db 592 HGTSSQPE 599

RESULT 15

PCT-US92-10621-3
 ; Sequence 3, Application PC/TUS9210621
 ; GENERAL INFORMATION:
 ; APPLICANT: Genentech, Inc.
 ; APPLICANT: Gorman, Cornelia M.,
 ; APPLICANT: Marriott, Dave,
 ; APPLICANT: Groskreutz Debraya J.
 ; TITLE OF INVENTION: Prohormone Convertase Transformed Cells and Polypeptide Syn
 ; NUMBER OF SEQUENCES: 54
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:

; NUMBER OF SEQUENCES: 19

APPLICATION NUMBER: US 08/344,639
FILING DATE: 14-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-688 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 298 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-852-344D-8

Query Match 69.2%; Score 27; DB 3; Length 298;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSTSSQ 6
Db 266 HSTSSE 271

RESULT 8
US-08-344-639E-8
Sequence 8, Application US/08344639E
Patent No. 6033668
GENERAL INFORMATION:
APPLICANT: Klein, Michel H
APPLICANT: Du, Run-Pan
APPLICANT: Ewasyshyn, Mary E
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERES
PROTECTION AGAINST PARAINFLUENZA VIRUS
TITLE OF INVENTION: AND RESPIRATORY SYNCYTIAL VIRUS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 330 University Avenue, 6th Floor
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/344,639E
FILING DATE: 14-NOV-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/001,554
FILING DATE: 06-JAN-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-391 MIS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163

TELEX: 065-24567 SIMBAS
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 298 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-344-639E-8

Query Match 69.2%; Score 27; DB 3; Length 298;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSTSSQ 6
Db 266 HSTSSE 271

RESULT 9
US-08-805-118-1
Sequence 1, Application US/08805118
Patent No. 5985604
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Bandman, Olga
TITLE OF INVENTION: NOVEL HUMAN SODIUM-DEPENDENT
TITLE OF INVENTION: PHOSPHATE CO-TRANSPORTER
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/805,118
FILING DATE: Filed Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0221 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 401 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAITUT02
CLONE: 754412
US-08-805-118-1

Query Match 69.2%; Score 27; DB 2; Length 401;
Best Local Similarity 75.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSTSSSQ 8

```
; Patent No. 5968776
; GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; APPLICANT: EWASYSHYN, Mary E
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,963C
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/838,189
; FILING DATE: 16-APR-1997
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-467-963C-8

Query Match 69.2%; Score 27; DB 2; Length 298;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HSTSSQ 6
Db 266 HSTSSE 271

RESULT 6
us-08-838-189D-8
; Sequence 8, Application US/08838189D
; Patent No. 5998169
; GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; APPLICANT: EWASYSHYN, Mary E
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,189D
; FILING DATE: 16-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-687 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-838-189D-8
```

```
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,189D
; FILING DATE: 16-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-687 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-838-189D-8

Query Match 69.2%; Score 27; DB 2; Length 298;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HSTSSQ 6
Db 266 HSTSSE 271

RESULT 7
us-08-852-344D-8
; Sequence 8, Application US/08852344D
; Patent No. 6017539
; GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; APPLICANT: EWASYSHYN, Mary E
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION
; TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY
; TITLE OF INVENTION: SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,344D
; FILING DATE: 07-MAY-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
```

Db 457 HSTSSQSQ 464

RESULT 2

PCT-US95-07085-2
; Sequence 2, Application PC/TUS9507085
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone Receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07085
; FILING DATE: 05-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-393
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-07085-2

Query Match 100.0%; Score 39; DB 4; Length 541;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSTSSQSQ 8

Db 457 HSTSSQSQ 464

RESULT 3

US-08-885-291-55
; Sequence 55, Application US/08885291A
; Patent No. 6057125
; GENERAL INFORMATION:
; APPLICANT: Takahashi, Joseph S.
; APPLICANT: Turek, Fred W.
; APPLICANT: Pinto, Lawrence H.
; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT
; FILE REFERENCE: 0290-5
; CURRENT APPLICATION NUMBER: US/08/885,291A
; CURRENT FILING DATE: 1997-06-30
; EARLIER APPLICATION NUMBER: 08/816,693
; EARLIER FILING DATE: 1997-03-13
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55

; LENGTH: 846
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-885-291-55

Query Match 71.8%; Score 28; DB 3; Length 846;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 STSSQSQ 8

Db 630 STSTQSQ 636

RESULT 4

US-09-107-847-2
; Sequence 2, Application US/09107847
; Patent No. 6100062
; GENERAL INFORMATION:
; APPLICANT: DUCKWORTH, DAVID
; APPLICANT: MICHALOVICH, DAVID
; TITLE OF INVENTION: NOVEL USE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P.O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,847
; FILING DATE: 30-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 97304996.8
; FILING DATE: 08-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-30003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 846 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-107-847-2

Query Match 71.8%; Score 28; DB 3; Length 846;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 STSSQSQ 8

Db 630 STSTQSQ 636

RESULT 5

US-08-467-963C-8
; Sequence 8, Application US/08467963C

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:49:21 ; Search time 97.15 Seconds
(without alignments)
1.380 Million cell updates/sec

Title: US-09-236-468A-2_COPY_457_464

Perfect score: 39

Sequence: 1 HSTSSQSQ 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 164575

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6.COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	100.0	541	3	US-08-468-011A-2
2	39	100.0	541	4	PCT-US95-07085-2
3	28	71.8	846	3	US-08-885-291-55
4	28	71.8	846	3	US-09-107-847-2
5	27	69.2	298	2	US-08-467-963C-8
6	27	69.2	298	2	US-08-838-189D-8
7	27	69.2	298	3	US-08-852-344D-8
8	27	69.2	298	3	US-08-344-639E-8
9	27	69.2	401	2	US-08-805-118-1
10	27	69.2	480	2	US-08-724-394A-11
11	27	69.2	681	5	5194595-19
12	27	69.2	732	1	US-08-481-626-2
13	27	69.2	732	1	US-08-712-241-2
14	27	69.2	753	1	US-08-712-241-3
15	27	69.2	753	4	PCT-US92-10621-3
16	27	69.2	753	4	PCT-US94-02233-3
17	27	69.2	1068	1	US-08-396-479B-12
18	27	69.2	1068	1	US-08-818-823-12
19	27	69.2	1075	4	PCT-US94-07297-41
20	27	69.2	1377	2	US-08-308-818-4
21	27	69.2	1400	1	US-08-080-255-7
22	27	69.2	1400	3	US-08-465-713-7
23	27	69.2	1400	4	PCT-US93-05857-7
24	27	69.2	1876	2	US-08-609-049A-12
25	27	69.2	1876	2	US-08-609-049A-28
26	26	66.7	117	1	US-07-752-101A-37
27	26	66.7	117	1	US-07-752-101A-42
28	26	66.7	234	3	US-08-836-236-4

29 26 66.7 234 3 US-08-836-236-5
30 26 66.7 264 2 US-08-960-022-20
31 26 66.7 440 1 US-08-483-140-27
32 26 66.7 440 2 US-08-485-938A-31
33 26 66.7 505 3 US-08-303-861-2
34 26 66.7 519 3 US-08-997-445D-2
35 26 66.7 532 4 PCT-US95-13975-2
36 26 66.7 551 1 US-08-120-960-2
37 26 66.7 694 2 US-08-701-240-2
38 26 66.7 695 2 US-08-701-240-4
39 26 66.7 947 2 US-08-887-518-2
40 26 66.7 947 2 US-09-023-321-2
41 26 66.7 947 2 US-09-032-475-2
42 26 66.7 1041 1 US-08-494-714-2
43 26 66.7 1041 4 PCT-US96-10782-2
44 26 66.7 1187 1 US-08-320-559-28
45 26 66.7 1187 3 US-08-545-860D-28

ALIGNMENTS

RESULT 1
US-08-468-011A-2
; Sequence 2, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-011A-2

Query Match 100.0%; Score 39; DB 3; Length 541;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
|||||||

XX Synthetic.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Modified-site 15
FT /note= "citrulline"
FT

XX WO9928344-A2.
FN 10-JUN-1999.
XX

XX 30-NOV-1998; 98WO-EP07714.
XX
XX 09-APR-1998; 98EP-0870078.
PR 28-NOV-1997; 97EP-0870195.
XX

XX (INNO-) INNOGENETICS NV.

XX Meheus L, Raymackers J, Union A;
PI
XX WPI; 1999-385357/32.
XX

XX New peptide derived from intermediate filament proteins

XX Claim 3; Page 51; 73pp; English.

XX Y22921-28 represent peptides derived from any variant of natural
CC flaggrin or any variant of intermediate filament proteins. The
CC peptides contain at least one citrulline residue which is crucial
CC for reacting with antibodies that are present in sera from patients
CC with rheumatoid arthritis. The peptides constitute immunogenic
CC determinants of antibodies present in patients with rheumatoid
CC arthritis. The peptides, antibodies, immunotoxins and intermediate
CC filament proteins can be used for the preparation of a therapeutic or
CC of a diagnostic for rheumatoid arthritis. The peptides can also be
CC used for identifying compounds which modulate the interaction between
CC an autoantigen and a rheumatoid arthritis specific autoantibody. The
CC products can also be used for the diagnosis and treatment of other
CC autoimmune diseases e.g. systemic lupus erythematosus, discoid lupus
CC erythematosus, scleroderma, dermatomyositis, or Sjogrens syndrome.
XX

SQ Sequence 17 AA;

Query Match 71.8%; Score 28; DB 20; Length 17;
Best Local Similarity 62.5%; Pred. No. 10;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 HSTSSQSQ 8
|||||
Db 1 hstsgegg 8

Search completed: November 8, 2000, 08:51:49
Job time: 261 sec

AC W26370;
XX 04-DEC-1997 (first entry)
XX Human steroid receptor coactivator-1 (SRC-1).
XX Steroid receptor coactivator-1; SRC-1; molecular switch;
KW gene therapy; transgenic animal.
XX Homo sapiens.
XX WO9710337-A1.
PN 20-MAR-1997.
XX 20-AUG-1996; 96WO-US13482.
XX 15-SEP-1995; 95US-0003784.
XX (BAYU) BAYLOR COLLEGE MEDICINE.
PA O'Walley BO, Onate S, Tsai M, Tsai SY;
PI WPI; 1997-202233/18.
DR N-PSDB; T84543.
XX Steroid receptor coactivator-1 polypeptide and DNA - regulated using
PT a molecular switch, used in gene therapy
XX Claim 3; Fig 1; 116pp; English.
XX This polypeptide comprises human steroid receptor coactivator-1
CC (SRC-1), a novel protein required for human progesterone receptor
CC (hPR) transactivational function. SRC-1 acts as a coactivator for
CC hPR by reversing receptor squelching. It is a coactivator for
CC multiple steroid receptor superfamily members. The truncated
CC C-terminal region of SRC-1 acts as a dominant negative regulator
CC of steroid receptor function. An SRC-1 related disease can be
CC treated by introducing SRC-1 nucleic acid (see T84543) into a host
CC cell and infusing the cells into the patient causing an increase in
CC the transcription of SRC-1 (claimed). A molecular switch can be
CC used to regulate expression of a nucleic acid cassette incorporating
CC an SRC-1 coding region for use in gene therapy. Transcription of a
CC target gene can be decreased by providing a nucleic acid encoding a
CC dominant-negative inhibitor of a SRC-1 polypeptide in a cell
CC containing the target gene (also claimed).
XX
SQ Sequence 1061 AA;

Query Match 74.4%; Score 29; DB 18; Length 1061;
Best Local Similarity 75.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
II III II
Db 51 hnsnsq 58

RESULT 14
W74319
ID W74319 standard; peptide; 12 AA.
XX W74319;
AC
XX 05-MAY-1999 (first entry)
DT HJ loop peptide K108H101.
DE
XX HJ loop; serine/threonine kinase; cancer; diabetes; obesity; therapy;
KW haemorrhagic shock; cardiovascular disease; immunosuppressive disorder;
KW inflammatory disorder; central nervous system disease; septic shock;
KW parkinson's disease; hypertension.

XX Synthetic.
OS
XX Key Location/Qualifiers
FH Modified-site 1
FT Modified-site /note= "Myristyl-Gly"
FT Modified-site 12
FT /note= "amidated"
XX WO9853051-A1.
PN 26-NOV-1998.
XX 20-MAY-1998; 98WO-US10321.
XX 21-MAY-1997; 97US-0861153.
XX (CHIL-) CHILDRENS MEDICAL CENT.
XX Ben-Sasson SA;
XX WPI; 1999-070143/06.
XX New peptide derivatives for modulating protein tyrosine kinase
PT activity - comprise a sequence corresponding to the HJ loop of a
PT protein tyrosine kinase, used for treating cancers or immune
PT disorders
XX Disclosure; Fig 7; 79pp; English.
PS
XX This sequence represents a peptide of the invention, and is a derivative
CC of the HJ loop of a serine/threonine kinase (STK). The peptides can be
CC used for the treatment of disorders caused by overactivity or
CC underactivity of a STK, e.g. cancer, diabetes, obesity, haemorrhagic
CC shock, cardiovascular diseases (e.g. thrombosis, atherosclerosis,
CC arteriosclerosis, cardiac hyper trophy, ischaemia, reperfusion injury and
CC hypertension), immunosuppressive and inflammatory disorders (e.g. asthma,
CC psoriasis, systemic lupus erythematosus, diabetes mellitus, suppression
CC of organ transplant rejection, multiple sclerosis, inflammatory bowel
CC disease and AIDS), central nervous system diseases (e.g. Alzheimer's
CC disease, stroke and trauma), septic shock, Parkinson's disease or
CC hypertension. The peptides can also be used to produce antibodies which
CC can be used to identify cells expressing the STK and to study the
CC intracellular distribution of the STK. In addition, the peptides can be
CC used to identify and quantitate ligands which bind the HJ loop of the STK
CC from which the peptide was derived.
XX
SQ Sequence 12 AA;

Query Match 71.8%; Score 28; DB 20; Length 12;
Best Local Similarity 85.7%; Pred. No. 6.9;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HSTSSQS 7
II IIII
Db 3 hsdssqs 9

RESULT 15
Y22924
ID Y22924 standard; peptide; 17 AA.
XX Y22924;
AC
XX 20-AUG-1999 (first entry)
DT Filagrin derived antigenic peptide.
DE
XX Filagrin; intermediate filament protein; antibody; rheumatoid arthritis;
KW antigen; immunotoxin; autoantigen; autoantibody; autoimmune disease;
KW systemic lupus erythematosus; discoid lupus erythematosus; scleroderma;
KW dermatomyositis; Sjogrens syndrome.

CC neurodegenerative disease, acute traumatic injury, fibrotic disease and
 CC autoimmune diseases (e.g. rheumatoid arthritis and sclerosis). The UNC-53
 CC polynucleotides can be used for recombinant production of the proteins,
 CC as a source of probes for detecting allelic variants and polymorphisms,
 CC for sequencing genomic DNA and for detecting UNC-53 expression; and as
 CC source of therapeutic antisense sequences. Cells that express the
 CC protein are used to identify regulators of cell shape, growth, motility
 CC and migration. They can also be used to identify proteins that are
 CC involved in signal transduction pathways also involving UNC-53, and to
 CC identify compounds that alter attachment of UNC-53 to microtubules. A
 CC target gene coupled to a UNC-53 encoding sequence may be used to deliver
 CC the target gene to a cellular microtubule or its plus ends. The present
 CC sequence represents the amino acid sequence of the C. briggsae UNC-53.

XX Sequence 1583 AA;

Query Match 76.9%; Score 30; DB 21; Length 1583;
 Best Local Similarity 85.7%; Pred. No. 4.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HSTSSQS 7
 |||||
 Db 423 hstssks 429

RESULT 11

Y70083
 ID Y70083 standard; Protein; 371 AA.

XX AC Y70083;

XX 05-JUN-2000 (first entry)

XX C. albicans vanadate resistance glycosylation 4 protein.

XX Vanadate resistance glycosylation 4; VRG4; nucleotide-sugar transporter;
 KW golgi-associated nucleotide-sugar; golgi GDP-mannose transporter;
 KW nucleotide-sugar transporter activity; inhibitor; antifungal;
 KW yeast growth.

XX Candida albicans.

XX WO200009550-A1.

XX 24-FEB-2000.

XX 12-AUG-1999; 99WO-US18402.

XX 14-AUG-1998; 98US-0096609.

XX 03-FEB-1999; 99US-0118370.

XX (UUNY) UNIV NEW YORK STATE RES FOUND.

XX Dean N;

XX WPI: 2000-224279/19.

XX N-PSDB; Z51062.

PT Measuring yeast nucleotide-sugar transporter activity useful for
 PT screening for inhibitors for use as antifungal agents -

XX Claim 49; Fig 7; 97pp; English.

CC The patent discloses methods for measuring yeast nucleotide-sugar
 CC transporter activity by determining an amount of golgi-associated
 CC nucleotide-sugar as an indicator of nucleotide-sugar transporter
 CC activity. This is used in screening for inhibitors of golgi nucleotide
 CC -sugar transporter activity for use as antifungal agents. The antifungal
 CC compound inhibits golgi GDP-mannose transport and is useful in a
 CC pharmaceutical composition for inhibiting growth of yeast in a patient.
 CC The methods, proteins, nucleotide sequences and antibodies are useful for
 CC detecting VRG4 and its gene in samples and for determining the efficacy

CC of treatment by an antifungal compound. The present sequence is
 CC Candida albicans vanadate resistance glycosylation 4 (VRG4)
 CC protein that transports GDP-mannose from the cytoplasm into the lumen of
 CC the golgi complex. VRG4 protein or immunogenic portions are useful in
 CC eliciting anti-VRG4 antibody to be used in diagnostic assay and as a
 CC therapeutic to inhibit transporter activity.

XX Sequence 371 AA;

Query Match 74.4%; Score 29; DB 21; Length 371;
 Best Local Similarity 85.7%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HSTSSQS 7
 |||||
 Db 38 hstssss 44

RESULT 12

W78487
 ID W78487 standard; Protein; 399 AA.

XX AC W78487;

XX 21-MAY-1999 (first entry)

XX Human SRC-1 ligand binding region.

XX Human; SRC-1; ligand binding; fusion protein; screening;
 KW intranuclear receptor protein.

XX Homo sapiens.

XX JPL1032767-A.

XX 09-FEB-1999.

XX 16-JUL-1997; 97JP-0191007.

XX 16-JUL-1997; 97JP-0191007.

XX (YAMA) YAMANOUCHI PHARM CO LTD.

XX WPI: 1999-183824/16.

XX N-PSDB; X21364.

XX Detection of an intranuclear receptor - using a recombinant
 PT intranuclear receptor-combining protein

XX Claim 3; Page 10-11; 23pp; Japanese.

XX This sequence represents the human SRC-1 protein ligand binding region
 CC corresponding to amino acids 384-782 of the full length protein. The
 CC protein can be fused to reporter proteins and used for screening for
 CC intranuclear receptor protein that bind the ligand.

XX Sequence 399 AA;

Query Match 74.4%; Score 29; DB 20; Length 399;
 Best Local Similarity 75.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
 |||||
 Db 48 hsnssnsq 55

RESULT 13

W26370
 ID W26370 standard; Protein; 1061 AA.

XX

PA (JANC) JANSSEN PHARM NV.
 XX Luyten WHML, De Raeymaeker MC, Geysen JJGH, Bogaert TAOE;
 PI Maerten LJS, Verhasselt P, Van De Craen M;
 XX WPI; 2000-116370/10.
 DR Novel proteins and nucleic acids e.g. for treating neurodegeneration -
 XX Disclosure; Fig 16; 146pp; English.
 XX The invention provides vertebrate (human) protein homologue of a UNC-53
 CC protein of Caenorhabditis elegans. The UNC-53 binds to microtubules or
 CC their plus ends. The UNC-53 sequences are used to promote neural
 CC regeneration, revascularization and wound healing; also for treating
 CC neurodegenerative disease, acute traumatic injury, fibrotic disease and
 CC autoimmune diseases (e.g. rheumatoid arthritis and sclerosis). The UNC-53
 CC polynucleotides can be used for recombinant production of the proteins,
 CC as a source of probes for detecting allelic variants and polymorphisms,
 CC for sequencing genomic DNA and for detecting UNC-53 expression; and as
 CC source of therapeutic antisense sequences. Cells that express the
 CC protein are used to identify regulators of cell shape, growth, motility
 CC and migration. They can also be used to identify proteins that are
 CC involved in signal transduction pathways also involving UNC-53, and to
 CC identify compounds that alter attachment of UNC-53 to microtubules. A
 CC target gene coupled to a UNC-53 encoding sequence may be used to deliver
 CC the target gene to a cellular microtubule or its plus ends. The present
 CC sequence represents the amino acid sequence of the C. elegans UNC-53
 CC splice variant (EST clone yk48066).
 XX
 SQ Sequence 1552 AA;
 Query Match 76.9%; Score 30; DB 21; Length 1552;
 Best Local Similarity 85.7%; Pred. No. 4.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 HSTSQS 7
 Db 313 hstssks 319
 |||||:
 RESULT 9
 W20057
 ID W20057 standard; Protein; 1583 AA.
 AC W20057;
 XX 10-SEP-1997 (first entry)
 DT C. elegans UNC-53 protein variant 7A.
 DE UNC-53; neuronal regeneration; revascularisation; wound healing;
 XX neurodegenerative disease; Alzheimer's disease; Huntington's;
 KW peripheral neuropathies; metastasis inhibition; cancer.
 KW Caenorhabditis elegans.
 XX OS WO9638555-A2.
 PN 05-DEC-1996.
 XX 31-MAY-1996; 96WO-EP02311.
 PF 31-MAY-1995; 95GB-0010944.
 XX (BOGA/) BOGAERT T.
 XX (STRI/) STRINGHAM E.
 PA (VAND/) VANDEKERCKHOVE J.
 XX Bogaert T, Stringham E, Vandekerckhove J;
 PI WPI; 1997-034369/03.
 DR

DR N-PSDB; T71315.
 XX Caenorhabditis elegans UNC-53 protein 8A and 7A variants - useful to
 PT promote neuronal regeneration, revascularisation or wound healing.
 PT Claim 22; Page 111-116; 278pp; English.
 PS UNC-53 protein variants 8A and 7A of Caenorhabditis elegans are new.
 XX The UNC-53 proteins and nucleic acids are useful as medicaments to
 CC promote neuronal regeneration, revascularisation or wound healing, or
 CC for treatment of chronic neurodegenerative diseases (e.g. Alzheimer's or
 CC Huntington's disease) or acute traumatic injuries. Transgenic cells and
 CC organisms transfected with UNC-53 cDNA can be used to determine whether
 CC a substance is an inhibitor or enhancer of the regulation of cell shape
 CC or motility or the direction of cell migration by screening for a
 CC phenotypic change in the cell. Inhibitors can be used to alleviate the
 CC spread of disease inducing cells or metastasis. Probes derived from the
 CC cDNA sequences can be used to identify homologues of the C. elegans
 CC unc-53 gene. The UNC-53 protein can be used to identify proteins which
 CC are active in the signal transduction pathway that can be used as
 CC mentioned above.
 XX
 SQ Sequence 1583 AA;
 Query Match 76.9%; Score 30; DB 18; Length 1583;
 Best Local Similarity 85.7%; Pred. No. 4.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 HSTSQS 7
 Db 422 hstssks 428
 |||||:
 RESULT 10
 Y85572
 ID Y85572 standard; Protein; 1583 AA.
 XX Y85572;
 AC 07-JUL-2000 (first entry)
 DT C. briggsae UNC-53 sequence.
 DE UNC-53; Caenorhabditis elegans; microtubule; neural regeneration;
 KW anticancer; anti-neurodegeneration; antifibrotic; anti-adhesive; human;
 KW antisclerotic; antimetastatic; anti-arthritis; autoimmune disease.
 XX Caenorhabditis briggsae.
 XX OS WO9963080-A1.
 PN 09-DEC-1999.
 XX 02-JUN-1999; 99WO-EP03848.
 PF 03-JUN-1998; 98GB-0011962.
 PR (JANC) JANSSEN PHARM NV.
 PA Luyten WHML, De Raeymaeker MC, Geysen JJGH, Bogaert TAOE;
 PI Maerten LJS, Verhasselt P, Van De Craen M;
 XX WPI; 2000-116370/10.
 DR Novel proteins and nucleic acids e.g. for treating neurodegeneration -
 XX Disclosure; Fig 3; 146pp; English.
 XX The invention provides vertebrate (human) protein homologue of a UNC-53
 CC protein of Caenorhabditis elegans. The UNC-53 binds to microtubules or
 CC their plus ends. The UNC-53 sequences are used to promote neural
 CC regeneration, revascularization and wound healing; also for treating
 CC neurodegenerative disease, acute traumatic injury, fibrotic disease and
 CC autoimmune diseases (e.g. rheumatoid arthritis and sclerosis). The UNC-53
 CC polynucleotides can be used for recombinant production of the proteins,
 CC as a source of probes for detecting allelic variants and polymorphisms,
 CC for sequencing genomic DNA and for detecting UNC-53 expression; and as
 CC source of therapeutic antisense sequences. Cells that express the
 CC protein are used to identify regulators of cell shape, growth, motility
 CC and migration. They can also be used to identify proteins that are
 CC involved in signal transduction pathways also involving UNC-53, and to
 CC identify compounds that alter attachment of UNC-53 to microtubules. A
 CC target gene coupled to a UNC-53 encoding sequence may be used to deliver
 CC the target gene to a cellular microtubule or its plus ends. The present
 CC sequence represents the amino acid sequence of the C. elegans UNC-53
 CC splice variant (EST clone yk48066).
 XX
 SQ Sequence 1552 AA;
 Query Match 76.9%; Score 30; DB 21; Length 1552;
 Best Local Similarity 85.7%; Pred. No. 4.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 HSTSQS 7
 Db 313 hstssks 319
 |||||:
 RESULT 9
 W20057
 ID W20057 standard; Protein; 1583 AA.
 AC W20057;
 XX 10-SEP-1997 (first entry)
 DT C. elegans UNC-53 protein variant 7A.
 DE UNC-53; neuronal regeneration; revascularisation; wound healing;
 XX neurodegenerative disease; Alzheimer's disease; Huntington's;
 KW peripheral neuropathies; metastasis inhibition; cancer.
 KW Caenorhabditis elegans.
 XX OS WO9638555-A2.
 PN 05-DEC-1996.
 XX 31-MAY-1996; 96WO-EP02311.
 PF 31-MAY-1995; 95GB-0010944.
 XX (BOGA/) BOGAERT T.
 XX (STRI/) STRINGHAM E.
 PA (VAND/) VANDEKERCKHOVE J.
 XX Bogaert T, Stringham E, Vandekerckhove J;
 PI WPI; 1997-034369/03.
 DR

DT 13-OCT-1999 (first entry)
 DE Human lung tumour protein SAL-25 2nd predicted amino acid sequence.
 XX
 KW Human; lung tumour protein; therapy; diagnosis; lung cancer; vaccine;
 KW immunotherapy; detection; inhibition.
 XX
 OS Homo sapiens.
 XX
 XX WO9938973-A2.
 XX
 XX PD 05-AUG-1999.
 XX
 XX PF 26-JAN-1999; 99WO-US01642.
 XX
 XX PR 22-DEC-1998; 98US-0219245.
 PR 28-JAN-1998; 98US-0015022.
 PR 28-JAN-1998; 98US-0015029.
 PR 18-MAR-1998; 98US-0040828.
 PR 18-MAR-1998; 98US-0040831.
 PR 23-JUL-1998; 98US-0122191.
 PR 23-JUL-1998; 98US-0122192.
 XX
 XX PA (CORI-) CORIXA CORP.
 XX
 XX PI Frudakis TN, Lodes MJ, Mohamath R, Reed SG;
 XX
 XX DR WPI: 1999-479187/40.
 DR N-PSDB; 207192.
 XX
 XX PT Lung tumour specific polynucleotides for inhibiting the development
 of lung cancer
 XX
 XX PS Claim 3; Page 148-149; 171pp; English.
 XX
 XX CC The present invention describes lung tumour specific polynucleotides
 and tumour antigens. 207144 to 207246 and 208301 to 208325 represent
 CC specifically claimed polynucleotides, and Y29486 to Y29571 represent
 CC amino acid sequences from the present invention. The lung tumour
 CC specific polynucleotides and polypeptides can be used in pharmaceutical
 CC compositions and vaccines to inhibit the development of lung cancer.
 CC They can also be used to detect lung cancer in a patient. Probes and
 CC antibodies derived from the lung tumour sequences are useful in
 CC detection of lung cancer.
 XX
 XX SQ Sequence 746 AA;

 Query Match 76.9%; Score 30; DB 20; Length 746;
 Best Local Similarity 75.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

 QY 1 HSTSSQSQ 8
 II :IIII
 Db 260 hsltsqsg 267

 RESULT 7
 W20056
 ID W20056 standard; Protein; 1528 AA.
 XX
 AC W20056;
 XX
 DT 10-SEP-1997 (first entry)
 XX
 DE C. elegans UNC-53 protein variant 8A.
 XX
 KW UNC-53; neuronal regeneration; revascularisation; wound healing;
 KW neurodegenerative disease; Alzheimer's disease; Huntington's;
 KW peripheral neuropathies; metastasis inhibition; cancer.
 XX
 OS Caenorhabditis elegans.
 XX

PN WO9638555-A2.
 XX
 PD 05-DEC-1996.
 XX
 PF 31-MAY-1996; 96WO-EP02311.
 XX
 PR 31-MAY-1995; 95GB-0010944.
 XX
 XX (BOGA/) BOGAERT T.
 PA (STRI/) STRINGHAM E.
 PA (VAND/) VANDEKERCKHOVE J.
 XX
 PI Bogaert T, Stringham E, Vandekerckhove J;
 XX
 DR WPI: 1997-034369/03.
 DR N-PSDB; T71314.
 XX
 PT Caenorhabditis elegans UNC-53 protein 8A and 7A variants - useful to
 promote neuronal regeneration, revascularisation or wound healing.
 XX
 XX PS Claim 20; Page 106-111; 278pp; English.
 XX
 CC UNC-53 protein variants 8A and 7A of Caenorhabditis elegans are new.
 CC The UNC-53 proteins and nucleic acids are useful as medicaments to
 CC promote neuronal regeneration, revascularisation or wound healing, or
 CC for treatment of chronic neurodegenerative diseases (e.g. Alzheimer's or
 CC Huntington's disease) or acute traumatic injuries. Transgenic cells and
 CC organisms transfected with UNC-53 cDNA can be used to determine whether
 CC a substance is an inhibitor or enhancer of the regulation of cell shape
 CC or motility or the direction of cell migration by screening for a
 CC phenotypic change in the cell. Inhibitors can be used to alleviate the
 CC spread of disease inducing cells or metastasis. Probes derived from the
 CC cDNA sequences can be used to identify homologues of the C. elegans
 CC unc-53 gene. The UNC-53 protein can be used to identify proteins which
 CC are active in the signal transduction pathway that can be used as
 CC mentioned above.
 XX
 XX SQ Sequence 1528 AA;

 Query Match 76.9%; Score 30; DB 18; Length 1528;
 Best Local Similarity 85.7%; Pred. No. 4.5e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 1 HSTSSQS 7
 IIIII:I
 Db 422 hstssks 428

 RESULT 8
 Y85577
 ID Y85577 standard; Protein; 1552 AA.
 XX
 AC Y85577;
 XX
 DT 07-JUL-2000 (first entry)
 XX
 DE C. elegans UNC-53 splice variant (EST clone yk480b6).
 XX
 KW UNC-53; Caenorhabditis elegans; microtubule; neural regeneration;
 KW anticancer; anti-neurodegeneration; antifibrotic; anti-adhesive; human;
 KW antisclerotic; antimetastatic; anti-arthritis; autoimmune disease.
 XX
 OS Caenorhabditis elegans.
 XX
 XX PN WO9963080-A1.
 XX
 PD 09-DEC-1999.
 XX
 PF 02-JUN-1999; 99WO-EP03848.
 XX
 XX 03-JUN-1998; 98GB-0011962.
 XX

CC cartilage, bone, nerve and muscle, particularly in cases of subglottic
 CC stenosis, chondromalacia patellae, osteoarthritis, joint surface
 CC lesions, neurodegeneration (e.g. Alzheimer's disease), myodegeneration or
 CC osteodegeneration. They also modulate Wnt-mediated signalling in cells,
 CC and are used to inhibit growth of Wnt-expressing tumours (particularly
 CC mammary or intestinal). The Frzb genes may also be used to identify
 CC specific modulators or as a growth factor for cells of the chondrocyte
 CC lineage in vitro, to stimulate wound healing, to promote angiogenesis,
 CC to prevent transplant rejection and as adjunct to chemotherapy or
 CC immunotherapy. Fragments of Frzb proteins are used for detecting genetic
 CC abnormalities associated with Frzb genes.

XX
 SQ Sequence 325 AA;

Query Match 76.9%; Score 30; DB 19; Length 325;
 Best Local Similarity 75.0%; Pred. No. 88;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
 |||l:lll
 Db 305 hsdstgsq 312

RESULT 4
 Y03233
 ID Y03233 standard; Protein; 325 AA.

XX AC Y03233;

XX DT 16-AUG-1999 (first entry)

XX DE Amino acid sequence of the bovine frezzled-like protein.

XX KW Bovine frezzled-like protein; FRZB; human frezzled-like protein;
 KW HFLP; frizzled protein family; differentiation-related disorder;
 KW agonist; antagonist; antibody; haematopoiesis; wound healing;
 KW cancer; inflammatory disorder; autoimmune disease;
 KW allergic reaction; ss.

XX OS Bos sp.

XX PN W09909152-A1.

XX PD 25-FEB-1999.

XX PF 11-AUG-1998; 98WO-US16701.

XX PR 10-APR-1998; 98US-0081438.

XX PR 12-AUG-1997; 97US-0055715.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Olsen HS, Ruben SM;

XX DR WPI; 1999-190159/16.

XX PT New isolated human frezzled-like protein - used to develop products
 PT for treating e.g. cancers, inflammatory and autoimmune diseases,
 PT infectious diseases and allergic reactions

XX PS Disclosure; Pages 137-138; 152pp; English.

XX CC This is the nucleotide sequence encoding the bovine frezzled
 CC protein (FRZB) used in the method of the invention involving
 CC the human frezzled-like protein (HFLP). It is a member of the
 CC frizzled protein family. It is used in the detection and treatment
 CC of differentiation-related disorders. In conditions where HFLP is
 CC under-expressed, its agonist is involved in the treatment.
 CC Antagonist and antibodies of HFLP are used in the treatment of
 CC disorder where HFLP is over-expressed, e.g. the regulation of
 CC haematopoiesis, and wound healing. HFLP products are used to
 CC develop products for treating e.g. cancers, inflammatory

CC and autoimmune diseases, infectious diseases and allergic
 CC reactions.
 XX
 SQ Sequence 325 AA;

Query Match 76.9%; Score 30; DB 20; Length 325;
 Best Local Similarity 75.0%; Pred. No. 88;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
 |||l:lll
 Db 305 hsdstgsq 312

RESULT 5
 P40254
 ID P40254 standard; Protein; 426 AA.

XX AC P40254;

XX DT 29-JUL-1992 (first entry)

XX DE Human IgD.

XX KW Immunoglobulin; B cell; autoimmune disease; rheumatism.

XX OS Homo sapiens.

XX PN W08404108-A.

XX PD 25-OCT-1984.

XX PF 24-AUG-1983; 83WO-JP00277.

XX PR 09-APR-1983; 83JP-0062563.

XX PA (KYOW) KYOWA HAKKO KOGYO.

XX PA (NABE/) NABESHIMA Y.

XX PI Nabeshima Y, Nabeshima Y, Ogata K;

XX DR WPI; 1984-276010/44.

XX DR N-PSDB; N40222.

XX PT Novel immunoglobulin coding DNA - prepd. by culturing IgD
 PT secretion cells, prodn. of cancer cells and column sepn.

XX PS Claim 3; Page 16; 24pp; Japanese.

XX CC IgD relates to maturation or differentiation of antibody-yielding
 CC B cells and autoimmune diseases such as rheumatism. IgD H chain-
 CC encoding DNA can be used as an agent for diagnosing diseases caused
 CC by IgD abnormality.

XX SQ Sequence 426 AA;

Query Match 76.9%; Score 30; DB 5; Length 426;
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HSTSSQSQ 8
 |||l:lll
 Db 265 hsdgsqsg 272

RESULT 6
 Y29512
 ID Y29512 standard; Protein; 746 AA.

XX AC Y29512;

XX

PT hyper-calcaemia, hypo- or hyper-phosphatemia, kidney stones, etc
 PS Claim 9; Fig 1A-E; 62pp; English.
 XX
 CC A novel 7-transmembrane receptor (W12695) has been identified as a
 CC human G-protein parathyroid hormone (PTH) receptor, designated
 CC H12697A. It shows 48.2% homology to the human PTH receptor. Its
 CC amino acid sequence was deduced from a cDNA clone (p59619) isolated
 CC from a human T cell lymphoma tissue cDNA library. Recombinant
 CC H12697A can be produced in transformed host cells and used to
 CC screen for (ant)agonist cpds. Agonists may be used to prevent or
 CC treat e.g. hypocalcaemia, hyperphosphatemia, hypoparathyroidism
 CC and chronic tetany by stimulating an increase in serum calcium
 CC levels. Antagonists can be used to inhibit the receptor e.g. for
 CC the treatment of osteoporosis, hypercalcaemia, hypoparathyroidism,
 CC hypophosphatemia, kidney stone, nephrolithiasis.
 XX
 SQ Sequence 541 AA;
 Query Match 100.0%; Score 39; DB 18; Length 541;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 HSTSSQSQ 8
 Db 457 hstsqsq 464
 RESULT 2
 ID P70535 standard; protein; 15 AA.
 AC
 AC P70535;
 XX
 XX 11-FEB-1991 (first entry)
 DT
 XX
 DE Antigenic oligopeptide of retrovirus oncogene fes.
 XX
 KW Retrovirus oncogene; tumour diagnosis; therapy.
 XX
 OS Homo sapiens.
 XX
 PN US4699877-A.
 XX
 XX 13-OCT-1987.
 PD
 XX 20-NOV-1984; 84US-0673469.
 PF
 XX 20-NOV-1984; 84US-0673469.
 PR
 XX (REGC) UNIV OF CALIFORNIA.
 PA
 XX Cline MJ, Slamon DJ;
 PI
 XX WPI; 1987-306540/43.
 DR
 XX
 PT Detecting human tumour - using probe specific for m-RNA
 PT complementary to a retro virus DNA capable of transforming normal
 PT cell to malignancy.
 XX
 XX Claim 19; Page 9; 10pp; English.
 PS
 XX Antibodies raised against this human oncogene fes expression
 CC product are used for identifying and treating malignant cells
 CC of fresh tumours in a human host.
 CC See also P70515-21, P70529-34 and P70536-40.
 CC
 XX Sequence 15 AA;
 SQ
 Query Match 76.98; Score 30; DB 8; Length 15;
 Best Local Similarity 75.0%; Pred. No. 3.5;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 HSTSSQSQ 8
 Db 2 hstsqsq 9
 RESULT 3
 ID W48694 standard; Protein; 325 AA.
 AC
 AC W48694;
 XX
 XX 04-SEP-1998 (first entry)
 DT
 XX Bovine growth-inducing protein Frzb sequence.
 DE
 XX
 KW Frzb; growth-inducing protein; bovine; human; Xenopus; cartilage;
 KW bone; nerve; muscle; tumour; Wnt-expressing tumour; myodegeneration;
 KW subglottic stenosis; chondromalacia patellae; osteoarthritis;
 KW joint surface lesion; neurodegeneration; Alzheimer's disease;
 KW osteodegeneration; angiogenesis; wound healing.
 XX
 OS Bos sp.
 XX
 FH Key Location/Qualifiers
 FT Peptide 7..27
 FT /note= "putative signal peptide"
 FT Cleavage-site 32..33
 FT /note= "putative signal peptide cleavage site"
 FT Protein 33..325
 FT /note= "mature peptide"
 FT Domain 35..147
 FT /note= "N-terminal domain"
 FT Region 73..94
 FT /note= "putative transmembrane region"
 XX
 PN W09816641-A1.
 XX
 XX 23-APR-1998.
 PD
 XX 08-OCT-1997; 97WO-US18362.
 PF
 XX 20-MAR-1997; 97US-0822333.
 PR
 XX 11-OCT-1996; 96US-0729452.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Hoang B, Luyten FP, Moos M, Wang S;
 PI
 XX WPI; 1998-251288/22.
 DR
 XX N-PSDB; V18253.
 XX
 XX New nucleic acid encoding human, bovine and Xenopus Frzb protein -
 PT and related proteins, antibodies, peptide(s), vectors and
 PT transformed cells, used to induce growth of cartilage, bone, nerve
 PT and muscle, also for inhibiting Wnt-expressing tumours
 XX
 XX Claim 4; Fig 1; 66pp; English.
 PS
 XX This represents a bovine growth-inducing protein Frzb. Bovine, human
 CC and Xenopus Frzb genes which are shown in V18253 to V18255 are related
 CC to the frizzled gene in Drosophila. The corresponding bovine, human and
 CC Xenopus Frzb protein sequences are shown in W48694 to W48696. A
 CC recombinant construct containing a Frzb encoding nucleic acid linked to
 CC a heterologous promoter in an expression vector can be used to produce
 CC recombinant Frzb proteins. The Frzb proteins may be formulated with
 CC fibrin glue, freeze-dried cartilage grafts or collagen (optionally also
 CC cartilage progenitor cells, chondroblasts or chondrocytes). They are
 CC coated on to, or mixed with, a (non-)resorbable matrix, or mixed with a
 CC biodegradable polymer. They modulate activity of the growth factors Wnt-1
 CC to 8. The Frzb proteins induce skeletal morphogenesis, embryonic pattern
 CC formation and tissue specificity and are used to induce growth of

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:51:48 ; Search time 138.73 Seconds
(without alignments)
1.972 Million cell updates/sec

Title: US-09-236-468A-2_COPY_457_464
Perfect score: 39
Sequence: 1 HSTSSQSQ 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues
Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36.*
1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	100.0	541	18 W12695	G-protein parathy
2	30	76.9	15	P70535	Antigenic oligopep
3	30	76.9	325	19 W48694	Bovine growth-indu
4	30	76.9	325	20 Y03233	Amino acid sequenc
5	30	76.9	426	5 P40254	Human IgD. Homo s
6	30	76.9	746	20 Y29512	Human lung tumour
7	30	76.9	1528	18 W20056	C. elegans UNC-53
8	30	76.9	1552	21 Y85577	C. elegans UNC-53
9	30	76.9	1583	18 W20057	C. elegans UNC-53
10	30	76.9	1583	21 Y85572	C. briggsae UNC-53
11	29	74.4	371	21 Y70083	C. albicans vanada
12	29	74.4	399	20 W78487	Human SRC-1 ligand

13	29	74.4	1061	18 W26370	Human steroid rece
14	28	71.8	12	20 W74319	HJ loop peptide K1
15	28	71.8	17	20 Y22924	Filagrin derived a
16	28	71.8	17	20 Y22944	Synthetic peptide
17	28	71.8	17	20 Y22949	Synthetic peptide
18	28	71.8	18	20 Y22938	Peptide derived fr
19	28	71.8	30	20 Y22926	Filagrin derived a
20	28	71.8	90	20 Y38406	Human secreted pro
21	28	71.8	330	20 Y22954	Human filagrin seq
22	28	71.8	330	20 Y22955	Human filagrin seq
23	28	71.8	330	20 Y22956	Human filagrin seq
24	28	71.8	330	21 Y82479	A. thaliana transc
25	28	71.8	335	21 Y82476	A. thaliana transc
26	28	71.8	846	19 W79533	Human CLOCK protei
27	28	71.8	846	20 W84565	Human HSCLOCK poly
28	28	71.8	846	21 Y32215	Human CLOCK protei
29	28	71.8	1395	20 Y13563	Drosophila Robo 1
30	28	71.8	1395	20 Y08401	Drosophila sp. ROB
31	27	69.2	15	20 W93564	RSV attachment G p
32	27	69.2	38	17 R91404	ME#2a and ME#2b 11
33	27	69.2	154	19 W72380	Arabidopsis pathog
34	27	69.2	232	20 W96314	Secreted G protein
35	27	69.2	258	11 R05671	Modified Pseudomon
36	27	69.2	258	11 R06614	Modified Pseudomon
37	27	69.2	258	11 R06615	Modified Pseudomon
38	27	69.2	258	11 R06616	Modified Pseudomon
39	27	69.2	258	11 R06617	Modified Pseudomon
40	27	69.2	258	11 R06618	Modified Pseudomon
41	27	69.2	258	11 R06619	Modified Pseudomon
42	27	69.2	258	11 R06620	Modified Pseudomon
43	27	69.2	258	11 R06686	Modified Pseudomon
44	27	69.2	258	15 R60880	Pseudomonas mendoc
45	27	69.2	258	16 R65238	Lipase with cutina

ALIGNMENTS

RESULT 1	
W12695	
ID	W12695 standard; Protein: 541 AA.
XX	
AC	W12695;
XX	
DT	31-MAY-1997 (first entry)
XX	
DE	G-protein parathyroid hormone receptor HLTDG74.
XX	
KW	G-protein parathyroid hormone receptor; HLTDG74; parathormone; PTH;
KW	calcium; signal transduction; agonist; antagonist; hypocalcaemia;
KW	hyperphosphataemia; hypoparathyroidism; chronic tetany;
KW	osteoporosis; hypercalcaemia; hypoparathyroidism; hypophosphataemia;
KW	kidney stone; nephrolithiasis; therapy; diagnosis.
OS	Homo sapiens.
XX	
PN	W09639433-Al.
XX	
PD	12-DEC-1996.
XX	
PF	05-JUN-1995; 95WO-US07085.
XX	
PR	05-JUN-1995; 95WO-US07085.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Li Y, Rosen CA, Ruben SM, Soppet DR;
XX	
DR	WPI; 1997-043068/04.
DR	N-PSDB; T59619.
XX	
PT	Human G-protein parathyroid hormone receptor, HLTDG74 - used to
PT	identify (ant)agonists, used in the treatment of hypo- or

OC Ephydroidea; Drosophilidae; Drosophila.

RP [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE; 20196006.
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclebb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE002760; AAF45469.1; -;
 DR FLYBASE; FBgn0033000; CG14464.
 SQ SEQUENCE 116 AA; 13715 MW; 64CD223B26E4FE93 CRC64;

Query Match 45.7%; Score 43; DB 5; Length 116;
 Best Local Similarity 53.8%; Pred. No. 15;
 Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 3 WKRTPPCGSRRCG 15

Db 74 WYNCRCGSTRCG 86

Search completed: November 8, 2000, 08:56:16
 Job time: 527 sec

Query Match 46.8%; Score 44; DB 11; Length 321;
Best Local Similarity 40.9%; Pred. No. 27;
Matches 9; Conservative 1; Mismatches 4; Indels 8; Gaps 1;

QY 2 DMKRTPP-----CGSRRCG 15
DB 6 DWVKHWPFGTDSRGCGRRRCG 27

RESULT 12
Q9SBX5 PRELIMINARY; PRT; 528 AA.
ID Q9SBX5;
AC Q9SBX5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE AUX1 PROTEIN.
GN AUX1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-NODAL TISSUE;
RA Wulff D., Hochholdinger F., Reuter K., Felix G.;
RT "Cloning of an AUX1 homologous CDNA from maize.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ011794; CAB65535.1; -;
DR INTERPRO: IPR002422; -;
DR PFAM: PF01490; Aa_trans; 1.
FT INIT_MET 0
SQ SEQUENCE 528 AA; 58937 MW; D10A6D2312691C74 CRC64;

Query Match 46.8%; Score 44; DB 10; Length 528;
Best Local Similarity 63.6%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 WKRTPPCGSRR 13
DB 494 WPRRPPCHRRR 504

RESULT 13
Q22024 PRELIMINARY; PRT; 543 AA.
ID Q22024;
AC Q22024; Q22023;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE EGL-43 PROTEIN.
GN EGL-43.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA Wilkinson J.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94150718.
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A., Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirsten J., Laister N., Latreille P., Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R., Smaclon N., Smith A., Sonnhammer E., Staden K., Sulston J., Thierri-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

elegans.";
RL Nature 368:32-38(1994).
DR EMBL: Z66515; CAA91353.1; -;
DR EMBL: Z66515; CAA91352.1; -;
DR HSSP: P07248; 2ADR.
DR INTERPRO: IPR000822; -;
DR PFAM: PF00096; zf-C2H2; 5.
DR PRINTS: PR00048; ZINC_FINGER.
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 4.
KW Zinc-finger; Metal-binding; DNA-binding.
SQ SEQUENCE 543 AA; 59356 MW; E3A287994095DA82 CRC64;

Query Match 46.8%; Score 44; DB 5; Length 543;
Best Local Similarity 63.6%; Pred. No. 42;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 RTPPCGSRRCG 15
DB 152 QTPPDGSHKCG 162

RESULT 14
Q26336 PRELIMINARY; PRT; 581 AA.
ID Q26336;
AC Q26336; Q26337;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE EGL-43.
GN EGL-43.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94040732.
RA Garriga G., Guenther C., Horvitz H.R.;
RT "Migrations of the Caenorhabditis elegans HSNs are regulated by egl-43, a gene encoding two zinc finger proteins.";
RL Genes Dev. 7:2097-2109(1993).
DR EMBL: S66757; AAB28819.1; -;
DR EMBL: S66936; AAB28820.1; -;
DR HSSP: P07248; 2ADR.
DR INTERPRO: IPR000822; -;
DR PFAM: PF00096; zf-C2H2; 6.
DR PRINTS: PR00048; ZINC_FINGER.
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 5.
KW Zinc-finger; Metal-binding; DNA-binding; Alternative splicing.
SQ SEQUENCE 581 AA; 63885 MW; 2B032B410681F0DD CRC64;

Query Match 46.8%; Score 44; DB 5; Length 581;
Best Local Similarity 63.6%; Pred. No. 45;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 RTPPCGSRRCG 15
DB 152 QTPPDGSHKCG 162

RESULT 15
Q9W5T6 PRELIMINARY; PRT; 116 AA.
ID Q9W5T6;
AC Q9W5T6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE CG14464 PROTEIN.
GN CG14464.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

```
RESULT 8
O60730 ID O60730 PRELIMINARY; PRT; 471 AA.
AC O60730;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE CDC14B3 PHOSPHATASE (EC 3.1.3.48).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Hao L., Baskerville C., Charbonneau H.;
RA Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Hao L., Baskerville C., Charbonneau H.;
RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DE EMBL: AF064105; AAC16662.2; -.
DR INTERPRO: IPR000340; -.
DR INTERPRO: IPR000387; -.
DR PFAM: PF00782; DSPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
KW Hydrolase.
SQ SEQUENCE 471 AA; 54175 MW; 7E55AF3DF2B39475 CRC64;

Query Match 47.3%; Score 44.5; DB 4; Length 471;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 3 WKRTPPCGSRRC 14
I III IIII
DB 11 WAAAPPC-SRRC 21

RESULT 9
O60729 ID O60729 PRELIMINARY; PRT; 498 AA.
AC O60729;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE CDC14B1 PHOSPHATASE (EC 3.1.3.48).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Hao L., Baskerville C., Charbonneau H.;
RA Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DE EMBL: AF064104; AAC16661.1; -.
DR INTERPRO: IPR000340; -.
DR INTERPRO: IPR000387; -.
DR PFAM: PF00782; DSPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
KW Hydrolase.
SQ SEQUENCE 498 AA; 56801 MW; CEE15EC4DC3B1DC7 CRC64;

Query Match 47.3%; Score 44.5; DB 4; Length 498;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
```

```
QY 3 WKRTPPCGSRRC 14
I III IIII
DB 11 WAAAPPC-SRRC 21

RESULT 10
Q92G15 ID Q92G15 PRELIMINARY; PRT; 4613 AA.
AC Q92G15;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE TYPE I POLYKETIDE SYNTHASE PIKAI.
GN PIKAI.
OS Streptomyces venezuelae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC15439;
RX MEDLINE; 98445333.
RA Xue Y., Zhao L., Liu H.W., Sherman D.H.;
RA "A gene cluster for macrolide antibiotic biosynthesis in Streptomyces
RT venezuelae: architecture of metabolic diversity.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:12111-12116(1998).
DR EMBL: AF079138; AAC69329.1; -.
DR HSSP; P25715; IMLA.
DR INTERPRO: IPR000169; -.
DR INTERPRO: IPR000255; -.
DR INTERPRO: IPR000794; -.
DR INTERPRO: IPR001227; -.
DR PFAM: PF00109; ketoacyl-synt; 3.
DR PFAM: PF00550; pp-binding; 3.
DR PFAM: PF00698; Acyl transf; 3.
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; 2.
DR PROSITE: PS00606; B_KETOACYL-SYNTHASE; 2.
DR PROSITE: PS00639; THIOL-PROTEASE_HIS; UNKNOWN_1.
DR PROSITE: PS50075; ACP_DOMAIN; 3.
KW Transferase.
SQ SEQUENCE 4613 AA; 477288 MW; A3BAF8D37CEC9383 CRC64;
```

```
Query Match 47.3%; Score 44.5; DB 2; Length 4613;
Best Local Similarity 56.2%; Pred. No. 2.3e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 1; Gaps 1;
```

```
QY 1 VDWKRTPPC-GSRRC 15
I IIII I IIII
DB 2006 IDWKRLPAEGSERTG 2021

RESULT 11
Q9QYA2 ID Q9QYA2 PRELIMINARY; PRT; 321 AA.
AC Q9QYA2;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE MITOCHONDRIAL OUTER MEMBRANE PROTEIN MOM35.
GN MOM35.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA Lee Rivera I., Schleiff E., Shore G.C.;
RA "Cloning and Characterization of a Mitochondrial Protein with High
RT Homology to Tom40.";
RL J. Bioenerg. Biomembr. 0:0-0(2000).
DR EMBL: AF109918; AAF21906.1; -.
SQ SEQUENCE 321 AA; 35251 MW; C90D5A730DAEFEB8D CRC64;
```

```

RT the Shope fibroma virus genome." ;
RL J. Virol. 50:408-416(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KASZA;
RX MEDLINE; 92074222.
RA Strayer D.S., Jerng H.H., O'Connor K.;
RT "Sequence and analysis of a portion of the genomes of Shope fibroma
RT virus and malignant rabbit fibroma virus that is important for viral
RT replication in lymphocytes." ;
RL Virology 185:585-595(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=KASZA;
RX MEDLINE; 20032074.
RA "The complete genome sequence of shope (Rabbit) fibroma virus." ;
RT Virology 264:319-343(1999).
DR EMBL; AF170722; AAF18020.1; -.
DR INTERPRO; IPR000210; -.
DR INTERPRO; IPR001798; -.
DR PFAM; PF00651; BTB. 1.
DR PEAM; PF01344; Kelch; 6.
SQ SEQUENCE 553 AA; 63121 MW; 321CCC1465C60B82 CRC64;

Query Match 50.0%; Score 47; DB 12; Length 553;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 DWKRTPPCGSRRRC 15
Db :||| ||| :|||
524 NWERIIPCKSPKCG 537

RESULT 5
O35130
ID O35130 PRELIMINARY; PRT; 244 AA.
AC O35130;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE BAC284H12 CHROMOSOME 6, COMPLETE SEQUENCE.
GN C2F.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA Ansari-Lari M.A., Oeltjen J.C., Schwartz S., Zhang Z., Muzny D.M.,
RA Lu J., Gorrell J.H., Chinault A.C., Belmont J.W., Miller W.,
RA Gibbs R.A.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Muzny D., Ansari-Lari M.A., Timms K.M., Yu W., Dugan S., Lu J.,
RA Shen Y., Rowland K., Liu W., Perez L., Ding Y., Haywood M., Jain A.,
RA Leal B., Logan O., Nguyen V., Savage L., Shen H., Worley K., Chen E.,
RA Forcum J., Aronson A.D., Chiu M.W., Gorrell J.H., Brundage E., Di W.,
RA Chinault C., Nelson D., Gibbs R.A.;
RA Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC002397; AAC36006.1; -.
SQ SEQUENCE 244 AA; 26974 MW; F9B1A6161D30F8BD CRC64;

```

```

Query Match 48.9%; Score 46; DB 11; Length 244;
Best Local Similarity 40.9%; Pred. No. 11;
Matches 9; Conservative 1; Mismatches 2; Indels 10; Gaps 1;

Qy 2 DWKRTPP-----CGSRR 13
Db ||| ||| ||| |||
21 DWETTPPKKRLGAGSKCGGRR 42

```

```

RESULT 6
Q9QR6
ID Q9QR6 PRELIMINARY; PRT; 407 AA.
AC Q9QR6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE GLYCOSYLTRANSFERASE.
GN MTMGIV.
OS Streptomyces argillaceus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC12956;
RA Fernandez E., Weissbach U., Sanchez Reillo C., Brana A.F., Mendez C.,
RA Rohr J., Salas J.A.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF077869; AAD5582.1; -.
DR INTERPRO; IPR000890; -.
DR PROSITE; PS01076; ACETATE_KINASE_2; UNKNOWN_1.
KW Transferase.
SQ SEQUENCE 407 AA; 43853 MW; 527A3F2BA370CA78 CRC64;

Query Match 47.9%; Score 45; DB 2; Length 407;
Best Local Similarity 61.5%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 DWKRTPPCGSRRRC 14
Db ||| ||| ||| |||
222 DWLRREPGRVC 234

RESULT 7
O43183
ID O43183 PRELIMINARY; PRT; 459 AA.
AC O43183;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE TYROSINE PHOSPHATASE (EC 3.1.3.48).
GN CDC14B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98037751.
RA Li L., Ernsting B.R., Wishart M.J., Lohse D.L., Dixon J.E.;
RT "A family of putative tumor suppressors is structurally and
RT functionally conserved in humans and yeast." ;
RL J. Biol. Chem. 272:29403-29406(1997).
DR EMBL; AF023158; AAB88293.1; -.
DR INTERPRO; IPR000340; -.
DR INTERPRO; IPR000387; -.
DR PFAM; PF00782; DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
KW Hydrolase.
SQ SEQUENCE 459 AA; 52751 MW; 7E5057E7677B2B61 CRC64;

Query Match 47.3%; Score 44.5; DB 4; Length 459;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 3 WKRTPPCGSRRRC 14
Db ||| ||| ||| |||
11 WAAAPPC-SRRRC 21

```

RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Poillard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Testor C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Welssenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003529; AAF49578.1; -.
DR FLYBASE: FBgn036535; CG6224.
DR INTERPRO: IPR00210; -.
DR INTERPRO: IPR001798; -.
DR PFAM: PF00651; BTB; 1.
DR PFAM: PF01344; Kelch; 6.
DR PRINTS: PR00501; KELCHREPEAT.
SQ SEQUENCE 623 AA; 68923 MW; DCD75E9F66BEC6B9 CRC64;

Query Match 54.38; Score 51; DB 5; Length 623;
Best Local Similarity 57.18; Pred. No. 4.3;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 DWKRTPPCGSRRCG 15
|||||
DB 348 DKWVAPMSKRRCG 361

RESULT 2
O61795
ID O61795 PRELIMINARY; PRT; 531 AA.
AC O61795;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE R12E2.1 PROTEIN.
GN R12E2.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodirinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE: 94150718.
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,
RA Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA Thiermy-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Goela D., Scheet P.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF067219; AAC17022.1; -.
DR INTERPRO: IPR000210; -.
RT "Tumorigenic poxviruses: construction of the composite physical map of

DR INTERPRO: IPR001798; -.
DR PFAM: PF00651; BTB; 1.
DR PFAM: PF01344; Kelch; 4.
DR PRINTS: PR00501; KELCHREPEAT.
SQ SEQUENCE 531 AA; 58647 MW; 07C91C5DD47A2C14 CRC64;

Query Match 52.18; Score 49; DB 5; Length 531;
Best Local Similarity 61.58; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 WKRTPPCGSRRCG 15
|||||
DB 335 WKCVAPMGKRRCG 347

RESULT 3
Q9WX19
ID Q9WX19 PRELIMINARY; PRT; 686 AA.
AC Q9WX19;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE FAMILY 19 CHITINASE (PRY1 ORF) PRECURSOR.
OS Aeromonas sp. 10S-24.
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonas group;
OC Aeromonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NO.10S-24;
RA Ueda M., Kawaguchi T., Miyatake K., Arai M.;
RT "Chitinolytic enzymes and their genes from *Aeromonas* sp No. 10S-24.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB025420; BAA76716.1; -.
DR INTERPRO: IPR000495; -.
DR INTERPRO: IPR000726; -.
DR INTERPRO: IPR002965; -.
DR PFAM: PF00182; Glyco_hydro_19; 1.
DR PRINTS: PR01217; PRICHEXTENS.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 686 FAMILY 19 CHITINASE (PRY1 ORF).
SQ SEQUENCE 686 AA; 72606 MW; 2A3F12F929B61DE9 CRC64;

Query Match 52.18; Score 49; DB 2; Length 686;
Best Local Similarity 66.78; Pred. No. 9.3;
Matches 10; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 3 WKRTPPCGSRRCG 15
|||||
DB 420 WTRTPPSRRSG 434

RESULT 4
Q9Q8U1
ID Q9Q8U1 PRELIMINARY; PRT; 553 AA.
AC Q9Q8U1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE GP14OR.
GN S14OR.
OS Rabbit fibroma virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Leporipoxvirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KASZA;
RX MEDLINE: 84165064.
RA Delange A.M., Macaulay C., Block W., Mueller T., McFadden G.;
RT "Tumorigenic poxviruses: construction of the composite physical map of

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:56:14 ; Search time 152.43 Seconds
(without alignments)
9.189 Million cell updates/sec

Title: US-09-236-468A-2_COPY_435_449

Perfect score: 94
Sequence: 1 VDMKRTPPGSRRCG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 297973

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL14.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	51	54.3	623	5 Q9VU05	Q9vuu5 drosophila
2	49	52.1	531	5 Q61795	O61795 caenorhabdi
3	49	52.1	686	2 Q9WXT9	Q9wxi9 aeromonas s
4	47	50.0	553	12 Q9Q8U1	Q9q8u1 rabbit fibr
5	46	48.9	244	11 Q35130	O35130 mus musculu
6	45	47.9	407	2 Q9RQR6	Q9rqr6 streptomyc
7	44.5	47.3	459	4 Q43183	O43183 homo sapien
8	44.5	47.3	471	4 Q60730	O60730 homo sapien
9	44.5	47.3	498	4 Q60729	O60729 homo sapien
10	44.5	47.3	4613	2 Q9ZG15	Q9zgi5 streptomyc
11	44	46.8	321	11 Q9QVA2	Q9qva2 mus musculu
12	44	46.8	528	10 Q9SBD5	Q9sbd5 zea mays (m
13	44	46.8	543	5 Q22024	Q22024 caenorhabdi
14	44	46.8	581	5 Q26336	Q26336 caenorhabdi
15	43	45.7	116	5 Q9W5N6	Q9w5n6 drosophila
16	43	45.7	575	13 Q9PWB7	Q9pwb7 brachydanio
17	42	44.7	74	2 Q9XAS0	Q9xas0 streptomyc
18	42	44.7	332	5 Q21451	Q21451 caenorhabdi
19	42	44.7	461	4 Q75072	Q75072 homo sapien

20	42	44.7	486	3 Q60106	O60106 schizosacch
21	42	44.7	537	3 Q9UST5	Q9ust5 schizosacch
22	42	44.7	604	3 Q13801	O13801 schizosacch
23	41.5	44.1	296	11 Q35171	Q35171 mus musculu
24	41.5	44.1	932	11 Q62030	Q62030 mus musculu
25	41.5	44.1	956	4 Q9Y4H1	Q9y4h1 homo sapien
26	41.5	44.1	956	4 Q9UEJ7	Q9uej7 homo sapien
27	41.5	44.1	969	4 Q9UEG7	Q9ueg7 homo sapien
28	41.5	44.1	3198	5 Q9U8G8	Q9u8g8 manduca sex
29	41	43.6	28	3 Q9UVM3	Q9uvm3 schizophyll
30	41	43.6	103	10 Q9SVR2	Q9svr2 urtica dioi
31	41	43.6	112	10 Q9ZFS1	Q9zfs1 urtica dioi
32	41	43.6	112	10 Q9SYR5	Q9syrs5 urtica dioi
33	41	43.6	112	10 Q9S7W3	Q9s7w3 urtica dioi
34	41	43.6	112	10 Q9S7K1	Q9s7k1 urtica dioi
35	41	43.6	112	10 Q9S7B3	Q9s7b3 urtica dioi
36	41	43.6	112	10 Q9S705	Q9s705 urtica dioi
37	41	43.6	113	10 Q9S7C2	Q9s7c2 urtica dioi
38	41	43.6	174	4 Q9Y5S9	Q9y5s9 homo sapien
39	41	43.6	282	4 Q9S765	Q9s765 homo sapien
40	41	43.6	372	10 Q9SVR1	Q9svr1 urtica dioi
41	40.5	43.1	161	2 Q82960	Q82960 burkholderi
42	40.5	43.1	240	2 Q9S1N4	Q9slt4 streptomyc
43	40.5	43.1	244	13 Q9PWL5	Q9pwl5 brachydanio
44	40.5	43.1	266	10 Q9SVI0	Q9svi0 arabidopsis
45	40.5	43.1	382	10 Q9SUM6	Q9sum6 arabidopsis

ALIGNMENTS

RESULT	1
Q9VU05	
ID	Q9VU05 PRELIMINARY; PRT; 623 AA.
AC	Q9VU05;
DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE	CG6224 PROTEIN.
GN	CG6224.
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=BERKELEY;
RX	MEDLINE; 20196006.
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.R., Yeandle M.D., Zhang Q., Chen L.X.,
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
RA	Wan K.H., Doyle C., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA	Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattier P.,
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferrier A., Fleischmann W.,
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

```
DR PFAM; PF00090; tsp_1; 6.  
KW Signal; Complement alternate pathway; Glycoprotein; Repeat.  
FT SIGNAL 1 26  
FT CHAIN 27 470  
FT DOMAIN 76 438  
FT REPEAT 76 133  
FT REPEAT 134 190  
FT REPEAT 191 254  
FT REPEAT 255 312  
FT REPEAT 313 376  
FT REPEAT 379 438  
FT CARBOHYD 429 429  
SQ SEQUENCE 470 AA; 51431 MW; FDC2B393DC7EC15F CRC64;  
  
Query Match 43.1%; Score 40.5; DB 1; Length 470;  
Best Local Similarity 50.0%; Pred. No. 46;  
Matches 8; Conservative 1; Mismatches 0; Indels 7; Gaps 1;  
  
Qy 6 TPFGS-----RRC 14  
|||||  
Db 393 TPFGANPTRVRQRC 408  
  
RESULT 14  
ID LCRP_PETMA STANDARD; PRT; 19 AA.  
AC Q10996;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE CORTICOSTATIN-RELATED PROTEIN LCRP.  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;  
OC Petromyzontiformes; Petromyzontidae; Petromyzon.  
RN [1]  
RP SEQUENCE.  
RX TISSUE=SKIN;  
RC MEDLINE; 96321324.  
RA Conlon J.M., Sower S.A.;  
RT "Isolation of a peptide structurally related to mammalian  
RT corticostatins from the lamprey Petromyzon marinus."  
RL Comp. Biochem. Physiol. 114B:133-137(1996).  
CC -!- FUNCTION: MAY HAVE MICROBICIDAL ACTIVITIES. MAY INHIBIT  
CC CORTICOTROPIN (ACTH) STIMULATED STEROIDOGENESIS AND THE MICROBIAL  
CC ACTIONS OF THE CORTICOSTATINS.  
CC -!- MASS SPECTROMETRY: MW=2201; MWERR=0.4; METHOD=ELECTROSPRAY.  
CC -!- SIMILARITY: BELONGS TO THE FAMILY OF CORTICOSTATIN/DEFENSIN  
CC PEPTIDES.  
KW Defensin; Antibiotic.  
FT DISULFID 1 18 BY SIMILARITY.  
FT DISULFID 3 9 BY SIMILARITY.  
FT DISULFID 8 17 BY SIMILARITY.  
SQ SEQUENCE 19 AA; 2209 MW; 8D9CEDC71A199AE5 CRC64;  
  
Query Match 42.6%; Score 40; DB 1; Length 19;  
Best Local Similarity 85.7%; Pred. No. 2.9;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 8 PCGSRRC 14  
||| |||  
Db 2 PCGRRRC 8  
  
RESULT 15  
ID VE7_HPVO5 STANDARD; PRT; 103 AA.  
AC P06932;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE E7 PROTEIN.  
  
GN E7.  
OS Human papillomavirus type 5.  
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Papillomavirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 87207670.  
RA Zachow K.R., Ostrow R.S., Faras A.J.;  
RT "Nucleotide sequence and genome organization of human papillomavirus  
RT type 5."  
RL Virology 158:251-254(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94086789.  
RA Deau M.C., Favre M., Jablonska S., Rueda L.A., Orth G.;  
RT "Genetic heterogeneity of oncogenic human papillomavirus type 5  
RT (HPV5) and phylogeny of HPV5 variants associated with  
RT epidermodysplasia verruciformis."  
RL J. Clin. Microbiol. 31:2918-2926(1993).  
CC -!- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING  
CC ACTIVITIES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; M17463; AAA46984.1; -;  
CC EMBL; X74618; CAA52690.1; -;  
CC PIR; G26277; W7WL5.  
CC INTERPRO; IPR00148; -;  
CC PFAM; PF00527; E7; 1.  
KW Early protein; Transcription regulation; Transforming protein;  
KW DNA-binding; Trans-acting factor.  
FT DOMAIN 58 61 C-XX-C MOTIF.  
FT DOMAIN 91 94 C-XX-C MOTIF.  
SQ SEQUENCE 103 AA; 11677 MW; A9FF74AD5C7EBBED CRC64;  
  
Query Match 42.6%; Score 40; DB 1; Length 103;  
Best Local Similarity 42.9%; Pred. No. 14;  
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
  
Qy 1 VDWKRTPPCGSRRC 14  
:| ||| | |  
Db 48 ISYKVIAPCGCRNC 61  
  
Search completed: November 8, 2000, 09:03:53  
Job time: 863 sec
```

```
DR PROSITE; PS00773; CHITINASE_19_1; 1.
DR PROSITE; PS00774; CHITINASE_19_2; 1.
KW Hydrolase; Glycosidase; Chitin degradation; Chitin-binding;
KW Lectin; Duplication; Fungicide; Signal.
FT SIGNAL 1 23
FT CHAIN 24 372 LECTIN/ENDOCHITININASE.
FT CHAIN 24 112 LECTIN.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT DOMAIN 24 69 CHITIN-BINDING 1 (BY SIMILARITY).
FT DOMAIN 70 112 CHITIN-BINDING 2 (BY SIMILARITY).
FT DOMAIN 113 128 SPACER.
FT DOMAIN 129 372 CHITINASE.
FT DISULFID 26 41 BY SIMILARITY.
FT DISULFID 35 47 BY SIMILARITY.
FT DISULFID 40 54 BY SIMILARITY.
FT DISULFID 58 62 BY SIMILARITY.
FT DISULFID 72 87 BY SIMILARITY.
FT DISULFID 81 93 BY SIMILARITY.
FT DISULFID 86 100 BY SIMILARITY.
FT DISULFID 105 109 BY SIMILARITY.
FT CARBOHYD 123 123 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 372 AA; 40541 MW; 28E9A2312BF08FD2 CRC64;

Query Match 43.6%; Score 41; DB 1; Length 372;
Best Local Similarity 75.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 PPCGSRRC 14
DB 79 PPCGQDRC 86

RESULT 12
STC_DROME
ID STC_DROME STANDARD; PRT; 1106 AA.
AC P40798;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SHUTTLE CRAFT PROTEIN.
GN STC.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Ovary;
RX MEDLINE; 96104568.
RA Stroumbakis N.D., Li Z., Tollas P.P.;
RT "A homolog of human transcription factor NF-X1 encoded by the
RT Drosophila shuttle craft gene is required in the embryonic central
RT nervous system.";
RL Mol. Cell. Biol. 16:192-201(1996).
CC -!- FUNCTION: PLAYS AN ESSENTIAL ROLE DURING THE LATE STAGES OF
CC EMBRYONIC NEUROGENESIS. MAY EITHER FINE-TUNE THE GUIDANCE OR THE
CC SPATIAL MAINTENANCE OF THE MIGRATING SNB AND IN NERVE ROOTS,
CC WHICH ARE COMPOSED OF AXONS ORIGINATING FROM DISTINCT GROUPS OF
CC MOTOR NEURONS AND MAY BE REQUIRED TO EITHER GUIDE OR MAINTAIN THE
CC POSITION OF THESE NERVES ALONG A DIRECT AND STRAIGHT PATH TO THEIR
CC ULTIMATE TARGETS IN PARTICULAR MUSCLE FIELDS. MAY PLAY A ROLE IN
CC EGG CHAMBER DEVELOPMENT AND/OR MAY CONFER ESSENTIAL MATERNAL
CC CONTRIBUTIONS TO THE EARLY EMBRYO.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- TISSUE SPECIFICITY: OVARIES AND EMBRYONIC CENTRAL NERVOUS SYSTEM.
CC -!- DEVELOPMENTAL STAGE: MAJOR EXPRESSION IS SEEN IN THE OVARIES WHILE
CC MODERATE LEVELS OF EXPRESSION ARE OBSERVED DURING EMBRYOGENESIS
CC AND THROUGHOUT SUBSEQUENT STAGES OF FLY DEVELOPMENT.
CC -!- SIMILARITY: TO YEAST YNL023C AND HUMAN NF1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U09306; AAB60255.1; -.
DR HSSP; P10968; IWGC.
DR TRANSFAC; T01688; -.
DR FLYBASE; FBgn0001978; stc.
DR INTERPRO; IPR000967; -.
DR INTERPRO; IPR001374; -.
DR PFAM; PF01424; R3H; 1.
DR PFAM; PF01422; zf-NF-X1; 8.
KW Transcription regulation; DNA-binding; Nuclear protein; Repeat;
KW Alternative splicing; RNA-binding.
FT DOMAIN 241 268 TANDDEM REPEATS OF R-D.
FT DOMAIN 464 904 7 X APPROXIMATE REPEATS, CYS-RICH.
FT REPEAT 464 500 1.
FT REPEAT 517 554 2.
FT REPEAT 575 614 3.
FT REPEAT 634 675 4.
FT REPEAT 723 762 5.
FT REPEAT 834 871 6.
FT REPEAT 865 904 7.
FT VARSPLIC 109 115 MISSING (IN SHORT ISOFORM).
SQ SEQUENCE 1106 AA; 123261 MW; 8CE146DB37BB8D65 CRC64;

Query Match 43.6%; Score 41; DB 1; Length 1106;
Best Local Similarity 43.8%; Pred. No. 86;
Matches 7; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 2 DWKRTPPCGSRR--CG 15
DB 820 DWKRTIPCGSQPNFSCG 835

RESULT 13
PROP_CAVPO
ID PROP_CAVPO STANDARD; PRT; 470 AA.
AC Q64181;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PROPERDIN PRECURSOR.
GN PPC.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SPLEEN;
RX MEDLINE; 96148617.
RA Maves K.K., Guenther S.T., Densen P., Moser D.R., Weiler J.M.;
RT "Cloning and characterization of the cDNA encoding guinea-pig
RT properdin: a comparison of properdin from three species.";
RL Immunology 86:475-479(1995).
CC -!- FUNCTION: A POSITIVE REGULATOR OF THE ALTERNATE PATHWAY OF
CC COMPLEMENT. IT BINDS TO AND STABILIZES THE C3-AND C5-CONVERTASE
CC ENZYME COMPLEXES.
CC -!- SIMILARITY: CONTAINS 6 TYPE-1 TSP REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; S81116; AAB35918.1; -.
DR INTERPRO; IPR000884; -.

```


FT VARSPLIC 472 969 PACE4.1).
 SQ SEQUENCE 969 AA; 106419 MW; A3599CC278D09B05 CRC64;

Query Match 44.1%; Score 41.5; DB 1; Length 969;
 Best Local Similarity 32.1%; Pred. No. 64;
 Matches 9; Conservative 0; Mismatches 4; Indels 15; Gaps 1;

QY 2 DWKRTPPCGS-----RRR 14
 ||| |||
 Db 874 DMKCVACGEGFYPEMPGLPKVCRRC 901

RESULT 10
 MF25_XENLA STANDARD; PRT; 287 AA.
 AC P16076;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE MYOGENIC FACTOR 25.
 GN MF25.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 90205830.
 RA Scales J.B., Olson E.N., Perry M.;
 RT "Two distinct Xenopus genes with homology to MyoD are expressed
 before somite formation in early embryogenesis.";
 RL Mol. Cell. Biol. 10:1516-1524(1990).
 CC -1- FUNCTION: MYOGENESIS.
 CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
 BHLH PROTEIN.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 TRANSCRIPTION FACTORS. "MYOGENIC FACTORS" SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M31118; AAA49902.1; -.
 DR PIR; C34783; C34783.
 DR HSSP; P10085; IMDY.
 DR INTERPRO; IPR001092; -.
 DR INTERPRO; IPR002545; -.
 DR INTERPRO; IPR003015; -.
 DR PFAM; PF01586; Basic; 1.
 DR PFAM; PF00010; HLH; 1.
 DR PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
 KW Myogenesis; Differentiation; Developmental protein; Nuclear protein;
 DNA-binding.
 FT DNA_BIND 96 108 BASIC DOMAIN.
 FT DOMAIN 109 148 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 SQ SEQUENCE 287 AA; 31896 MW; B2A73231B8CB25A6 CRC64;

Query Match 43.6%; Score 41; DB 1; Length 287;
 Best Local Similarity 58.3%; Pred. No. 25;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 DWKRTPPCGSRR 13
 | :||| |||
 Db 180 DGMNSPPCSRR 191

RESULT 11
 AGI_URTDI STANDARD; PRT; 372 AA.
 AC P11218;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE LECTIN/ENDOCHITINASE PRECURSOR (EC 3.2.1.14) (AGGLUTININ) (UDA).
 GN UDA1.
 OS Urtica dioica (Great nettle) (Stinging nettle).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids 1;
 OC Rosales; Urticaceae; Urtica.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 92283806.
 RA Lerner D.R., Raikhel N.V.;
 RT "The gene for stinging nettle lectin (Urtica dioica agglutinin)
 encodes both a lectin and a chitinase.";
 RL J. Biol. Chem. 267:11085-11091(1992).
 RN [2]
 RP ERRATUM.
 RX MEDLINE; 93054574.
 RA Lerner D.R., Raikhel N.V.;
 RL J. Biol. Chem. 267:22694-22694(1992).
 RN [3]
 RP SEQUENCE OF 24-112.
 RX MEDLINE; 92183920.
 RA Beintema J.J., Peumans W.J.;
 RT "The primary structure of stinging nettle (Urtica dioica) agglutinin.
 A two-domain member of the hevein family.";
 RL FEBS Lett. 299:131-134(1992).
 RN [4]
 RP PRELIMINARY SEQUENCE OF 24-54.
 RA Chapot M.-P., Feumans W.J., Strosberg A.D.;
 RT "Extensive homologies between lectins from non-leguminous plants.";
 RL FEBS Lett. 195:231-234(1986).
 CC -1- FUNCTION: THIS PROTEIN FUNCTIONS BOTH AS A CHITINASE AND AS A
 N-ACETYL-D-GLUCOSAMINE BINDING LECTIN. IT INHIBITS THE GROWTH OF
 SEVERAL PHYTOPATHOGENIC CHITIN-CONTAINING FUNGI. POSSESSES ALSO
 INSECTICIDAL ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES OF
 N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.
 CC -1- TISSUE SPECIFICITY: RHIZOMES AND INFLORESCENCE WITH IMMATURE
 SEEDS.
 CC -1- PTM: DO UNDERGO A PROCESSING EVENT TO RELEASE A VERY SMALL PROTEIN
 (8.5 KDA, 86 AA) CONTAINING ONLY THE TWO CHITIN-BINDING DOMAINS.
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS 1A WHICH IS SIMILAR TO
 CLASS 1B IN THE CATALYTIC PART BUT WHICH INCLUDE A N-TERMINAL
 CHITIN-BINDING LECTIN DOMAIN (BELONGS TO FAMILY 19 OF GLYCOSYL
 HYDROLASES).
 CC -1- SIMILARITY: CONTAINS TWO COPIES OF A CHITIN-BINDING DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M87302; AAA34219.1; -.
 DR PIR; B23616; B23616.
 DR PIR; S20437; S20437.
 DR PIR; A44298; A44298.
 DR HSSP; P27275; IMMC.
 DR INTERPRO; IPR000726; -.
 DR INTERPRO; IPR001002; -.
 DR PFAM; PF00182; Glyco_hydro_19; 1.
 DR PFAM; PF00187; chitin_binding; 2.
 DR PRINTS; PR00451; CHITINBINDING.
 DR PROSITE; PS00026; CHITIN_BINDING; 2.

```

SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-HYPOTHALAMUS; AND PITUITARY;
RX MEDLINE; 94349873.
RT Johnson R.C., Darlington D.N., Hand T.A., Bloomquist B.T., Mains R.E.;
RA "PAC64: a subtilisin-like endoprotease prevalent in the anterior
RT pituitary and regulated by thyroid status.";
RL Endocrinology 135:1178-1185(1994).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: HIGH EXPRESSION IN THE ANTERIOR PITUITARY AND
CC IN SEVERAL BRAIN REGIONS, THE ATRIUM, AND THE VENTRICLE.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILASE FAMILY. HIGH SIMILARITY WITH OTHER FURIN-LIKE ENZYMES.
-----
THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL; L31894; AAA61987.1; -.
DR HSPSP; Q99405; 1MPT.
DR INTERPRO; IPR000209; -.
DR INTERPRO; IPR002884; -.
DR PFAM; PF01483; P; 1.
DR PFAM; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR KW Hydrolyase; Serine protease; Glycoprotein; Zymogen; Signal; Repeat.
FT SIGNAL 1 45
FT PROPEP 46 132
FT CHAIN 133 937
FT DOMAIN 663 937
FT ACT_SITE 186 186
FT ACT_SITE 227 227
FT ACT_SITE 401 401
FT SITE 534 536
FT DOMAIN 661 909
FT REPEAT 661 707
FT REPEAT 708 758
FT REPEAT 759 806
FT REPEAT 807 855
FT REPEAT 856 909
FT CARBOHYD 240 240
FT CARBOHYD 882 882
FT CARBOHYD 900 900
FT SEQUENCE 937 AA; 104053 MW; F386557C33705C8 CRC64;
Query Match 44.1%; Score 41.5; DB 1; Length 937;
Best Local Similarity 32.1%; Pred. No. 62;
Matches 9; Conservative 0; Mismatches 4; Indels 15; Gaps 1;
QY 2 DWKRTPPCGS-----RRC 14
DB 842 DWKCVACGEGFVPEMPGLPHKVCRCR 869
RESULT 9
PAC4_HUMAN STANDARD; PRT; 969 AA.
AC P29122;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SUBTILISIN-LIKE PROTEASE PAC4 PRECURSOR (EC 3.4.21.-).
GN PAC4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

RP	SEQUENCE FROM N.A.
RP	MEDLINE; 92075167.
RA	Kiefer M.C., Tucker J.E., Joh R., Landsberg K.E., Saltman D.,
RA	Barr P.J.;
RT	"Identification of a second human subtilisin-like protease gene in
RT	the tes/lps region of chromosome 15.";
RL	DNA Cell Biol. 10:757-769(1991).
RP	(2)
RP	SEQUENCE FROM N.A. (PACE4B).
RP	MEDLINE; 98021085.
RA	Tsuji A., Hine C., Tamai Y., Yonemoto K., Mori K., Yoshida S.,
RA	Bando M., Sakai E., Mori K., Akamatsu T., Matsuda Y.;
RT	"Genomic organization and alternative splicing of human PACE4 (SPC4),
RT	kinin-like processing endopeptidase.";
RL	J. Biochem. 122:438-452(1997).
CC	-1- SUBCELLULAR LOCATION: SECRETED.
CC	-1- ALTERNATIVE PRODUCTS: TWO VARIANTS SEEM TO BE PRODUCED BY
CC	ALTERNATIVE SPLICING OF A SINGLE GENE. THEY DIFFER AFTER
CC	LYS-471.
CC	-1- TISSUE SPECIFICITY: WIDESPREAD, WITH COMPARATIVELY HIGHER LEVELS
CC	IN THE LIVER. PACE4.1 WAS ONLY FOUND IN THE LIVER.
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC	SUBTILASE FAMILY. HIGH SIMILARITY WITH OTHER FURIN-LIKE ENZYMES.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@lsb-sib.ch).
CC	-----
DR	EMBL; M80482; AAS9998.1; .
DR	EMBL; AB001914; BAA21620.1; .
DR	EMBL; AB001898; BAA21620.1; JOINED.
DR	EMBL; AB001900; BAA21620.1; JOINED.
DR	EMBL; AB001901; BAA21620.1; JOINED.
DR	EMBL; AB001902; BAA21620.1; JOINED.
DR	EMBL; AB001903; BAA21620.1; JOINED.
DR	EMBL; AB001904; BAA21620.1; JOINED.
DR	EMBL; AB001905; BAA21620.1; JOINED.
DR	PIR; A39490; A39490.
DR	HSSP; Q99405; IMPT.
DR	MM; 167405; .
DR	INTERPRO: IPR000209; .
DR	INTERPRO: IPR002884; .
DR	PFAM; PF01483; P; 1.
DR	PFAM; PF00082; Peptidase_s8; 1.
DR	PRINTS; PR00723; SUBTILISIN.
DR	PROSITE; PS00136; SUBTILASE ASP; 1.
DR	PROSITE; PS00137; SUBTILASE_HIS; 1.
DR	PROSITE; PS00138; SUBTILASE_SER; 1.
KW	Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;
KW	Alternative splicing; Repeat.
FT	SIGNAL 1 63 POTENTIAL.
FT	PROPEP 64 149 POTENTIAL.
FT	CHAIN 150 969 SUBTILISIN-LIKE PROTEASE PACE4.
FT	DOMAIN 695 969 CYS-RICH REGION.
FT	ACT_SITE 205 205 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 246 246 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 420 420 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	SITE 553 555 CELL ATTACHMENT SITE (POTENTIAL).
FT	DOMAIN 686 941 5 X TANDEM REPEATS, CYS-RICH.
FT	REPEAT 686 739 1.
FT	REPEAT 740 790 2.
FT	REPEAT 791 838 3.
FT	REPEAT 839 887 4.
FT	REPEAT 888 941 5.
FT	CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 914 914 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 932 932 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLIC 471 471 K -> KGAAVAFWWTIGWPNV (IN ISOFORM

```

Query Match      44.1%; Score 41.5; DB 1; Length 937;
Best Local Similarity 32.1%; Pred. No. 62;
Matches          9; Conservative 0; Mismatches 4; Indels 15; Gaps 1;

QY      2 DWKRTTPCGS-----RRC 14
      |||  |  ||
      |||  |  ||
Db       842 DWKCVACGGCFYPEEMPGLPHKVCRC 869
      |||  |  ||
      |||  |  ||

RESULT 9
PAC4_HUMAN
ID PAC4_HUMAN      STANDARD;      PRT;      969 AA.
AC P29132;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SUBTILISIN-LIKE PROTEASE PACE4 PRECURSOR (EC 3.4.21.-).
GN PACE4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

KW Transcription regulation; DNA-binding.
 FT DNA_BIND 95 107 BASIC DOMAIN.
 FT DOMAIN 108 147 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 FT CONFLICT 255 256 LQ -> A (IN REF. 2).
 SQ SEQUENCE 289 AA; 32306 MW; DOA985DC66136284 CRC64;

Query Match 45.8%; Score 44; DB 1; Length 289;
 Best Local Similarity 87.5%; Pred. No. 9;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 6 TPFGSRR 13
 :|||||
 Db 186 SPFGSRR 193

RESULT 6
 UL47_HSV1F STANDARD; PRT; 664 AA.
 AC P08313;
 DT 01-AUG-1998 (Rel. 08, Created)
 DT 01-AUG-1998 (Rel. 08, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last annotation update)
 DE VIRION PROTEIN UL47 (82/81 KDA TEGUMENT PROTEIN) (VMW82/81) (VP13/14).
 CN UL47.
 OS Herpes simplex virus (type 1 / strain F).
 OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 (1)
 RP SEQUENCE FROM N.A.
 RA McKnight J.L.C., Pellett P.E., Jenkins F.J., Roizman B.;
 RX MEDLINE: 87141362.
 RT "Characterization and nucleotide sequence of two herpes simplex virus
 RT 1 genes whose products modulate alpha-trans-inducing factor-dependent
 RT activation of alpha genes."
 RL J. Virol. 61:992-1001(1987).
 CC -1- FUNCTION: MODULATOR OF ALPHA-TIF (VMW65 PHOSPHOPROTEIN) TRANS-
 CC ACTIVATION. POSSIBLY UL47 MAY HAVE KINASE ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: MAJOR TEGUMENT PROTEIN OF THE VIRIONS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE LATER STAGES OF INFECTION.
 CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL47,
 CC EHV-1 13, AND VZV 11.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M15621; AAA45767.1; -
 CC PIR; A26133; TNBE70.
 DR Transcription regulation; Trans-acting factor; Structural protein;
 KW Late protein.
 SQ SEQUENCE 664 AA; 70526 MW; 05A1C56B1B73B4EA CRC64;

Query Match 45.7%; Score 43; DB 1; Length 664;
 Best Local Similarity 85.7%; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 WKRTPPC 9
 :|||||
 Db 571 WARTPPC 577

RESULT 7
 BAC1_MOUSE STANDARD; PRT; 739 AA.
 AC P97302;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE TRANSCRIPTION REGULATOR PROTEIN BACH1 (BTB AND CNC HOMOLOG 1).
 GN BACH1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 (1)
 RN SEQUENCE FROM N.A.
 RP STRAIN-BALB/C;
 RC MEDLINE: 97042438.
 RA Oyake T., Itoh K., Motohashi H., Hayashi N., Hoshino H., Nishizawa M.,
 RA Yamamoto M., Igarashi K.;
 RT "Bach proteins belong to a novel family of BTB-basic leucine zipper
 RT transcription factors that interact with MafK and regulate
 RT transcription through the NF-E2 site."
 RL Mol. Cell. Biol. 16:6083-6095(1996).
 CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS AS REPRESSOR OR
 CC ACTIVATOR. BINDS, IN-VITRO, TO NF-E2 BINDING SITES. PLAY IMPORTANT
 CC ROLES IN COORDINATING TRANSCRIPTION ACTIVATION AND REPRESSION BY
 CC MAFK.
 CC -1- SUBUNIT: HETERODIMER OF BACH1 AND MAFK.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS.
 CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY. CNC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 BTB DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D86603; BAA13137.1; -
 CC HSPG; P34707; LSKN.
 DR MGD; MGI:894680; BACH1.
 DR INTERPRO; IPR000210; -
 DR INTERPRO; IPR001871; -
 DR PFAM; PF00651; BTB; 1.
 DR PFAM; PF00170; BZIP; 1.
 DR PROSITE; PS00097; BTB; 1.
 DR PROSITE; PS00036; BZIP_BASIC; 1.
 KW Transcription regulation; Activator; Repressor; DNA-binding;
 KW Nuclear protein.
 FT DOMAIN 34 100 BTB
 FT DNA_BIND 565 580 BASIC MOTIF.
 FT DOMAIN 588 610 LEUCINE-ZIPPER.
 SQ SEQUENCE 739 AA; 81373 MW; CE2DB606B05F6E32 CRC64;

Query Match 45.7%; Score 43; DB 1; Length 739;
 Best Local Similarity 70.0%; Pred. No. 30;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 5 RTPPGSRR 14
 :|||
 Db 180 QTPQDSRR 189

RESULT 8
 PAC4_RAT STANDARD; PRT; 937 AA.
 ID PAC4_RAT
 AC O63415;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE SUBTILISIN-LIKE PROTEASE PACE4 PRECURSOR (EC 3.4.21.-).
 GN PACE4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 (1)

```
CC or send an email to license@isb-sib.ch).
DR EMBL; AF034632; AAC26081.1; -.
DR GCRDB; GCR_2494; -.
DR MIM; 602885; -.
DR INTERPRO; IPR000276; -.
DR PFAM; PF00001; 7tm_1; 2.
DR PRINTS; PR00237; GPCRHHODPSN.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 35 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 36 56 1 (POTENTIAL).
FT DOMAIN 57 74 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 75 94 2 (POTENTIAL).
FT DOMAIN 95 112 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 113 134 3 (POTENTIAL).
FT DOMAIN 135 157 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 158 178 4 (POTENTIAL).
FT DOMAIN 179 246 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 247 270 5 (POTENTIAL).
FT DOMAIN 271 298 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 299 320 6 (POTENTIAL).
FT DOMAIN 321 334 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 335 358 7 (POTENTIAL).
FT DOMAIN 359 412 CYTOPLASMIC (POTENTIAL).
FT DISULFID 111 235 BY SIMILARITY.
FT CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 412 AA; 45344 MW; C13FF6165012DEF3 CRC64;

Query Match
Best Local Similarity 52.1%; Score 49; DB 1; Length 412;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 WKRTPPCGSRRC 14
+ + + + +
Db 19 WPALPCDERRC 30

RESULT 4
VG46_BPPF1
ID VG46_BPPF1 STANDARD; PRT; 424 AA.
AC P25131;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DE 46.5 KDA PROTEIN (ORF 424).
OS Bacteriophage Pfl.
OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 25102-B1;
RX MEDLINE; 91186399.
RA Hill D.F., Short N.J., Perham R.N., Petersen G.B.;
RT "DNA sequence of the filamentous bacteriophage Pfl.";
RL J. Mol. Biol. 218:349-364(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52107; CAA36334.1; -.
DR PIR; S15146; S15146.
DR PIR; S20702; S20702.
SQ SEQUENCE 424 AA; 46539 MW; 108EA88EA407BEC7 CRC64;

Query Match
Best Local Similarity 48.9%; Score 46; DB 1; Length 424;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
Best Local Similarity 50.0%; Pred. No. 6.4;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 DWKRTPPCGSRRC 15
+ + + + +
Db 370 DWKQVTTCAGROAG 383

RESULT 5
MYOD_XENLA
ID MYOD_XENLA STANDARD; PRT; 289 AA.
AC P13904;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE MYOBLAST DETERMINATION PROTEIN 1 HOMOLOG (MYOGENIC FACTOR 1).
GN MYOD OR MFL.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90059936.
RA Hopwood N.D., Pluck A., Gurdon J.B.;
RT "MyoD expression in the forming somites is an early response to
RT mesoderm induction in Xenopus embryos.";
RL EMBO J. 8:3409-3417(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90205830.
RA Scales J.B., Olson E.N., Perry M.;
RT "Two distinct xenopus genes with homology to MyoD1 are expressed
RT before somite formation in early embryogenesis.";
RL Mol. Cell. Biol. 10:1516-1524(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90353188.
RA Harvey R.P.;
RT "The Xenopus MyoD gene: an unlocalised maternal mRNA predates
RT lineage-restricted expression in the early embryo.";
RL Development 108:669-680(1990).
CC -----
CC -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC BHLH PROTEIN.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- DEVELOPMENTAL STAGE: EXPRESSION IS SPECIFIC TO THE MESODERM OF THE
CC GASTRULA.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS. "MYOGENIC FACTORS" SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X16106; CAA34232.1; -.
DR EMBL; M31116; AAA49900.1; -.
DR PIR; S06952; S06952.
DR PIR; A34783; A34783.
DR PIR; A60099; A60099.
DR HSSP; P10085; IMDY.
DR TRANSFAC; T00524; -.
DR INTERPRO; IPR001092; -.
DR INTERPRO; IPR002546; -.
DR INTERPRO; IPR003015; -.
DR PFAM; PF01586; Basic; 1.
DR PFAM; PF00010; HLH; 1.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
KW Myogenesis; Differentiation; Developmental protein; Nuclear protein;
```

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 24
FT CHAIN 25
FT DOMAIN 27
FT TRANSMEM 146
FT TRANSMEM 146
FT TRANSMEM 170
FT TRANSMEM 177
FT TRANSMEM 197
FT TRANSMEM 238
FT TRANSMEM 261
FT TRANSMEM 276
FT TRANSMEM 298
FT TRANSMEM 317
FT TRANSMEM 338
FT TRANSMEM 365
FT TRANSMEM 384
FT TRANSMEM 395
FT TRANSMEM 418
FT TRANSMEM 511
FT CARBOHYD 106
FT CARBOHYD 116
FT CARBOHYD 121
SQ SEQUENCE 550 AA; 62235 MW; 2ADD14DBA68A9BF8 CRC64;

Query Match 100.0%; Score 94; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 6e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDWKRTPPCGSRRRC 15
Db 435 VDWKRTPPCGSRRRC 449

RESULT 2
PTH2_RAT
ID PTH2_RAT STANDARD; PRT; 546 AA.
AC P70555;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PARATHYROID HORMONE RECEPTOR PRECURSOR (PTH2 RECEPTOR).
GN PTHR2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1]
RX SEQUENCE FROM N.A.
RA MEDLINE; 96426194.
RA Usdin T.B., Bonner T.I., Harta G., Mezey E.;
RT "Distribution of parathyroid hormone-2 receptor messenger ribonucleic acid in rat.";
RL Endocrinology 137:4285-4297(1996).
CC -1- FUNCTION: THIS IS A SPECIFIC RECEPTOR FOR PARATHYROID HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYL CYCLASE. PTHR2 MAY BE RESPONSIBLE FOR PTH EFFECTS IN A NUMBER OF PHYSIOLOGICAL SYSTEMS. IT MAY PLAY A SIGNIFICANT ROLE IN PANCREATIC FUNCTION. PTHR2 PRESENCE IN NEURONS INDICATES THAT IT MAY FUNCTION AS A NEUROTRANSMITTER RECEPTOR.
CC -1- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN BRAIN, ARTERIAL AND CARDIAC ENDOTHELIUM. FOUND AS WELL IN SPERM, IN THE HEAD OF THE EPIDIDYMIS. LOWER EXPRESSION IS FOUND IN VASCULAR SMOOTH MUSCLE, EXOCRINE PANCREAS, TESTIS AND PLACENTA.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

EMBL; U55836; AAC52849.1; --
DR GCRDB; GCR_1413; --
DR INTERPRO; IPR000832; --
DR PFAM; PF00002; 7tm.2; 1.
DR PRINTS; PR00249; GPCRSECRETIN
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 24
FT CHAIN 25
FT DOMAIN 27
FT TRANSMEM 144
FT TRANSMEM 168
FT TRANSMEM 175
FT TRANSMEM 195
FT TRANSMEM 236
FT TRANSMEM 259
FT TRANSMEM 274
FT TRANSMEM 296
FT TRANSMEM 314
FT TRANSMEM 335
FT TRANSMEM 362
FT TRANSMEM 381
FT TRANSMEM 392
FT TRANSMEM 415
FT CARBOHYD 51
FT CARBOHYD 106
FT CARBOHYD 116
FT CARBOHYD 121
SQ SEQUENCE 546 AA; 61800 MW; 2825AE4040313527 CRC64;

Query Match 69.1%; Score 65; DB 1; Length 546;
Best Local Similarity 60.0%; Pred. No. 0.012;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VDWKRTPPCGSRRRC 15
Db 432 IDWKKAPCGGHRYG 446

RESULT 3
GP38_HUMAN
ID GP38_HUMAN STANDARD; PRT; 412 AA.
AC O43193;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PUTATIVE G PROTEIN-COUPLED RECEPTOR GPR38.
GN GPR38.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
[1]
RX SEQUENCE FROM N.A.
RA MEDLINE; 98110578.
RA McKee K.K., Tan C.P., palyha O.C., Liu J., Feighner S.D.,
RA Hreniuk D.L., Smith R.G., Howard A.D., van der Ploeg L.H.T.;
RT "Cloning and characterization of two human G protein-coupled receptor genes (GPR38 and GPR39) related to the growth hormone secretagogue and neurotensin receptors.";
RL Genomics 46:426-434(1997).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN THYROID, STOMACH, AND BONE MARROW.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 09:03:52 ; Search time 58.45 Seconds
(without alignments)
8.199 Million cell updates/sec

Title: US-09-236-468A-2_COPY_435_449
Perfect score: 94
Sequence: 1 VDWKRTPPCGSRRCG 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 87993

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	94	100.0	550	1 PTR2_HUMAN	P49190 homo sapien
2	65	69.1	546	1 PTH2_RAT	P70555 rattus norv
3	49	52.1	412	1 GP38_HUMAN	O43193 homo sapien
4	46	48.9	424	1 V646_BPPF1	P25131 bacterioph
5	44	46.8	289	1 MYOD_XENLA	P13904 xenopus lae
6	43	45.7	664	1 U147_HSV1F	P88313 herpes simp
7	43	45.7	739	1 BAC1_MOUSE	P97302 mus musculu
8	41.5	44.1	937	1 PAC4_RAT	Q63415 rattus norv
9	41.5	44.1	969	1 PAC4_HUMAN	P29122 homo sapien
10	41	43.6	287	1 ME25_XENLA	P16076 xenopus lae
11	41	43.6	372	1 AC1_URTDI	P11218 urtica dioi
12	41	43.6	1106	1 STC_DROME	P40798 drosophila
13	40.5	43.1	470	1 PROP_CAVPO	Q64181 cavia porce
14	40	42.6	19	1 LCRP_PETMA	Q10996 petromyzon
15	40	42.6	103	1 VE7_HPV05	P06932 human papil
16	40	42.6	103	1 VE7_HPV5B	P26559 human papil
17	40	42.6	398	1 OPRM_MOUSE	P42866 mus musculu
18	40	42.6	514	1 VLL_HPV47	P22424 human papil
19	40	42.6	607	1 COG_V_HUMAN	P51512 homo sapien
20	40	42.6	607	1 COG_V_RAT	P35548 rattus norv
21	40	42.6	649	1 SCAG_HUMAN	P51170 homo sapien
22	40	42.6	697	1 KPCG_MOUSE	P05697 mus musculu
23	40	42.6	697	1 KPCG_RABIT	P10829 oryctolagus
24	40	42.6	1234	1 CFAH_MOUSE	P6909 mus musculu
25	39.5	42.0	360	1 CCPH_HSV5A	Q10106 herpesvirus
26	39.5	42.0	853	1 NUG2_RHIME	P56914 rhizobium m
27	39.5	42.0	918	1 N1A_CUCMA	P17569 cucurbita m
28	39	41.5	398	1 OPRM_RAT	P33535 rattus norv
29	39	41.5	697	1 KPCG_HUMAN	P05129 homo sapien
30	39	41.5	831	1 SVFB_MYCTU	P94985 mycobacteri
31	38.5	41.0	217	1 SGS3_DROSI	P13729 drosophila
32	38.5	41.0	307	1 SGS3_DROME	P02840 drosophila
33	38.5	41.0	404	1 K1MB_HUMAN	Q14525 homo sapien

34	38.5	41.0	411	1 YE1A_ECOLI	P25889 escherichia
35	38	40.4	151	1 HS11_WHEAT	P12810 triticum ae
36	38	40.4	153	1 HS11_SOYBN	P02519 glycine max
37	38	40.4	153	1 HS13_SOYBN	P04793 glycine max
38	38	40.4	154	1 HS14_SOYBN	P04794 glycine max
39	38	40.4	158	1 HS11_PEA	P19243 pismus sativ
40	38	40.4	158	1 HS12_MEDSA	P27880 medicago sa
41	38	40.4	161	1 HS13_ARATH	P19037 arabidopsis
42	38	40.4	161	1 HS16_SOYBN	P05478 glycine max
43	38	40.4	164	1 CDN1_HUMAN	P38936 homo sapien
44	38	40.4	166	1 Y516_SYNY3	Q55837 synecocyst
45	38	40.4	192	1 HS41_SOYBN	P20236 glycine max

ALIGNMENTS

RESULT 1

PTR2_HUMAN

ID PTR2_HUMAN STANDARD; PRT; 550 AA.

AC P49190;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE PARATHYROID HORMONE RECEPTOR PRECURSOR (PTH2 RECEPTOR).

GN PTHR2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

CC [1]

RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.

RC TISSUE-BRAIN;

RX MEDLINE; 95318121

RA Usdin T.B., Gruber C., Bonner T.I.;

RT "Identification and functional expression of a receptor selectively

recognizing parathyroid hormone, the PTH2 receptor.,"

J. Biol. Chem. 270:15455-15458(1995).

RN [2]

RP SEQUENCE OF 26-40 AND 306-550 FROM N.A.

RX MEDLINE; 97079671.

RA Usdin T.B., Modi W., Bonner T.I.;

RT "Assignment of the human PTH2 receptor gene (PTH2) to chromosome 2q33

by fluorescence in situ hybridization.,"

Genomics 37:140-141(1996).

CC -!- FUNCTION: THIS IS A SPECIFIC RECEPTOR FOR PARATHYROID HORMONE. THE

ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE

ADENYL CYCLASE.

-!- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN BRAIN AND PANCREAS.

CC ALSO EXPRESSED IN THE TESTIS.

-!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U25128; AAC50157.1; -

DR EMBL; U47124; AAA96796.1; -

DR EMBL; U47129; AAC50767.1; -

DR EMBL; U47125; AAC50767.1; JOINED.

DR EMBL; U47126; AAC50767.1; JOINED.

DR EMBL; U47127; AAC50767.1; JOINED.

DR EMBL; U47128; AAC50767.1; JOINED.

DR GCRDB; GCR.2003; -

DR MIN; 601469; -

DR INTERPRO; IPR000832; -

DR PFAM; PF00002; 7tm_2; 1.

DR PRINTS; PR00249; GPCRSECRETIN.

DR PROSITE; PS00649; G-PROTEIN_RECEP_F2.1; 1.

DR PROSITE; PS00650; G-PROTEIN_RECEP_F2.2; 1.

R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Churcher, C.M.
submitted to the EMBL Data Library, June 1998

A;Reference number: Z21855
A;Accession: T39456
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-486 <LYN>
A;Cross-references: EMBL:AL023780; NID:el295812; PIDN:CAA19328.1; GSPDB:GN00067; SPDB:SH
A;Experimental source: strain 972h-; cosmid c14F5
C;Genetics:
A;Gene: SPDB:SPBC14F5.10c
A;Map position: 2
F;165-212/Domain: RING finger homology <RRN>

Query Match 44.7%; Score 42; DB 2; Length 486;
Best Local Similarity 53.8%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 DWKRTPPCGSRRC 14
|| | ||| |
Db 27 DWPVTLP CGGTVC 39

RESULT 13

T40247
hypothetical protein SPBC336.09c - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T40247
R;Borzym, K.; Beck, A.; Reinhardt, R.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, October 1999
A;Reference number: Z21916
A;Accession: T40247
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-537 <BOR>
A;Cross-references: EMBL:AL121815; PIDN:CAB58161.1; GSPDB:GN00067; SPDB:SPBC336.09c
A;Experimental source: strain 972h-; cosmid c336
C;Genetics:
A;Gene: SPDB:SPBC336.09c
A;Map position: 2
A;Introns: 37/3

Query Match 44.7%; Score 42; DB 2; Length 537;
Best Local Similarity 46.2%; Pred. No. 82;
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 DWKRTPPCGSRRC 14
: | ||| | |
Db 4 NWFGPPCSVRNC 16

RESULT 14

T37870
RNA-binding / Ran zinc finger proteinrotein - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T37870
R;Skellton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z21751
A;Accession: T37870
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-604 <SKE>
A;Cross-references: EMBL:Z98597; PIDN:CAB11213.1; GSPDB:GN00066; SPDB:SPAC17H9.04c
A;Experimental source: strain 972h-; cosmid c17H9
C;Genetics:
A;Gene: SPDB:SPAC17H9.04c
A;Map position: 1

Query Match 44.7%; Score 42; DB 2; Length 604;
Best Local Similarity 57.1%; Pred. No. 90;
Matches 8; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

Qy 2 DWKRTPPCGSRRC 15
||| ||| ||
Db 448 DWK----CGSRGCG 457

RESULT 15

I52527
PACE4A - mouse (fragment)

C;Species: Mus musculus (house mouse)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 16-Jul-1999
C;Accession: I52527
R;Hosaka, M.; Murakami, K.; Nakayama, K.
Biomed. Res. 15, 383-390, 1994
A;Title: PACE4A is a ubiquitous endoprotease that has similar but not identical subst
A;Reference number: I52527
A;Accession: I52527
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-932 <RES>
A;Cross-references: GB:D50060; NID:g769700; PIDN:BAA08777.1; PID:g769701
C;Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
F;172-410/Domain: subtilisin homology <SBT>

Query Match 44.1%; Score 41.5; DB 2; Length 932;
Best Local Similarity 32.1%; Pred. No. 1.5e+02;
Matches 9; Conservative 0; Mismatches 4; Indels 15; Gaps 1;

Qy 2 DWKRTPPCGS-----RRC 14
||| | ||
Db 837 DWKCVACGEGFYPEMPGLPHKVCRC 864

Search completed: November 8, 2000, 08:53:32
Job time: 363 sec

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24250
R:Wilkinson, J.
submitted to the EMBL Data Library, October 1995
A:Reference number: Z19863
A:Accession: T24250
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-543 <WIL>
A:Cross-references: EMBL:Z66515; PIDN:CAA91353.1; GSPDB:GN00020; CESP:R53.3a
A:Experimental source: clone R53
C:Genetics:
A:Gene: CESP:R53.3a
A:Map position: 2
A:Introns: 40/3; 74/3; 118/2; 172/2; 210/3; 272/3; 305/2; 473/3; 497/3

Query Match 46.8%; Score 44; DB 2; Length 543;
Best Local Similarity 63.6%; Pred. No. 43;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 RTPPCGSRRCG 15
Db 152 QTTPDGSHKCG 162

RESULT 8

A49073
HSN motor neurons migration regulator (alternatively spliced) Egl-43 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 20-Mar-1998
C:Accession: A49073
R:Gariga, G.; Guenther, C.; Horvitz, H.R.
Genes Dev. 7, 2097-2109, 1993
A:Title: Migrations of the Caenorhabditis elegans HSNs are regulated by egl-43, a gene expressed in HSN neurons
A:Reference number: A49073; MUID:94040732
A:Accession: A49073
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-581 <GAK>
A:Cross-references: GB:S66757; NID:g452950; PID:g452951
A:Note: sequence extracted from NCBI backbone (NCBIN:139438, NCBIPI:139439)

Query Match 46.8%; Score 44; DB 2; Length 581;
Best Local Similarity 63.6%; Pred. No. 45;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 RTPPCGSRRCG 15
Db 152 QTTPDGSHKCG 162

RESULT 9

TNBE70
70.5K alpha trans-inducing protein - human herpesvirus 1 (strain F)
C:Species: human herpesvirus 1
C:Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999
C:Accession: A26133
R:McKnight, J.L.C.; Pellett, P.E.; Jenkins, F.J.; Roizman, B.
J. Virol. 61, 992-1001, 1987
A:Title: Characterization and nucleotide sequence of two herpes simplex virus 1 genes which encode proteins that induce transcription of the alpha genes
A:Reference number: A93031; MUID:87141362
A:Accession: A26133
A:Molecule type: DNA
A:Residues: 1-664 <MCK>
A:Cross-references: GB:M15621; NID:g330056; PIDN:AAA45767.1; PID:g330057
C:Comment: This protein may reduce transcriptional activity of the alpha genes.
C:Superfamily: herpesvirus 70.5K alpha trans-inducing protein
C:Keywords: trans-inducing protein; transcription regulation

Query Match 45.7%; Score 43; DB 1; Length 664;
Best Local Similarity 85.7%; Pred. No. 70;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 WKRTPPC 9
Db 571 WARTPPC 577

RESULT 10

T34608
hypothetical protein SC10A7.05 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C:Accession: T34608
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, S.M.; et al.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z21549
A:Accession: T34608
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-74 <SAU>
A:Cross-references: EMBL:AL078618; PIDN:CAB44515.1; GSPDB:GN00070; SCOEDB:SC10A7.05
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC10A7.05

Query Match 44.7%; Score 42; DB 2; Length 74;
Best Local Similarity 77.8%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 TPPCGSRRC 14
Db 4 TPRCASRRC 12

RESULT 11

T23653
hypothetical protein M01F1.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T23653
R:Sims, M.
submitted to the EMBL Data Library, October 1994
A:Reference number: Z19778
A:Accession: T23653
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-332 <WIL>
A:Cross-references: EMBL:Z46381; PIDN:CAA86514.1; GSPDB:GN00021; CESP:M01F1.8
A:Experimental source: clone M01F1
C:Genetics:
A:Gene: CESP:M01F1.8
A:Map position: 3
A:Introns: 54/3; 123/3; 188/3; 289/1

Query Match 44.7%; Score 42; DB 2; Length 332;
Best Local Similarity 52.9%; Pred. No. 56;
Matches 9; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

Qy 1 VDWRKT--PPCGSRRCG 15
Db 106 VDWKYTKNPPTLAKVCG 122

RESULT 12

T39456
zinc finger protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jun-2000
C:Accession: T39456


```
C:Species: Streptomyces venezuelae
C>Date: 05-Oct-1999 #sequence_revision 05-Oct-1999 #text_change 05-Oct-1999
C:Accession: T17409
R:Xue, Y.; Zhao, L.; Liu, H.W.; Sherman, D.H.
Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998
A:Title: A gene cluster for macrolide antibiotic biosynthesis in streptomyces venezue
A:Reference number: Z18773; MUID:98445333
A:Accession: T17409
A>Status: preliminary
A:Residues: 1-4613 <XUE>
A:Cross-references: EMBL:AF079138; NID:g38008326; PID:g3800834; PIDN:AAC69329.1

Query Match 47.3%; Score 44.5; DB 3; Length 4613;
Best Local Similarity 56.2%; Pred. No. 2e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 VDWKRTPPC-GSRRCG 15
:||||| | |||||
Db 2006 IDWKRLFAAEGSRTG 2021
:||||| | |||||

RESULT 6
A34783
myogenesis protein MyoDa - African clawed frog
N:Alternate names: myoblast determination protein; myogenic factor 1
C:Species: Xenopus laevis (African clawed frog)
C>Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 09-Sep-1997
C:Accession: A34783; S06952; A60099
R:Scales, J.B.; Olson, E.N.; Perry, M.
Mol. Cell. Biol. 10, 1516-1524, 1990
A:Title: Two distinct Xenopus genes with homology to MyoD1 are expressed before somit
A:Reference number: A34783; MUID:90205830
A:Accession: A34783
A:Molecule type: mRNA
A:Residues: 1-289 <SCA>
A:Cross-references: GB:M31116
R:Hopwood, N.D.; Pluck, A.; Gurdon, J.B.
EMBO J. 8, 3409-3417, 1989
A:Title: MyoD expression in the forming somites is an early response to mesoderm indu
A:Reference number: S06952; MUID:90059936
A:Accession: S06952
A:Molecule type: mRNA
A:Residues: 1-289 <HOP>
A:Cross-references: EMBL:X16106; NID:g64906; PID:g64907
R:Harvey, R.P.
Development 108, 669-680, 1990
A:Title: The Xenopus MyoD gene: an unlocalised maternal mRNA predates lineage-restric
A:Reference number: A60099; MUID:90353188
A:Accession: A60099
A:Molecule type: mRNA
A:Residues: 1-289 <HAR>
C:Genetics:
A:Gene: mfi; MyoDa
C:Function:
A:Description: activates expression of skeletal muscle-specific genes
A:Pathway: myogenesis; skeletal muscle differentiation
A>Note: expressed exclusively in skeletal muscle and first activated during gastrulat
C:Keywords: alternative splicing; differentiation; DNA binding; skeletal muscle; tran

Query Match 46.8%; Score 44; DB 2; Length 289;
Best Local Similarity 87.5%; Pred. No. 26;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 TPPCGSR 13
:|||||
Db 186 SPPCGSR 193

RESULT 7
T24250
hypothetical protein R53.3a - Caenorhabditis elegans

C:Species: Streptomyces venezuelae
C>Date: 05-Oct-1999 #sequence_revision 05-Oct-1999 #text_change 05-Oct-1999
C:Accession: T17409
R:Xue, Y.; Zhao, L.; Liu, H.W.; Sherman, D.H.
Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998
A:Title: A gene cluster for macrolide antibiotic biosynthesis in streptomyces venezue
A:Reference number: Z18773; MUID:98445333
A:Accession: T17409
A>Status: preliminary
A:Residues: 1-4613 <XUE>
A:Cross-references: EMBL:AF079138; NID:g38008326; PID:g3800834; PIDN:AAC69329.1

Query Match 53.2%; Score 50; DB 2; Length 446;
Best Local Similarity 39.1%; Pred. No. 5;
Matches 9; Conservative 1; Mismatches 3; Indels 10; Gaps 1;

QY 3 WKRTPPCGS-----RRCG 15
:|||||
Db 331 WSQTPPCSGASVPVKRTACRRCG 353
:|||||

RESULT 3
T33099
hypothetical protein R12E2.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T33099
R:Goela, D.; Scheet, P.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid R12E2.
A:Reference number: Z21281
A:Accession: T33099
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-531 <GOE>
A:Cross-references: EMBL:AF067219; PIDN:AAC17022.1; GSPDB:GN00019; CESP:R12E2.1
A:Experimental source: strain Bristol N2; clone R12E2
C:Genetics:
A:Gene: CESP:R12E2.1
A:Map position: 1
A:Introns: 34/3; 60/3; 95/1; 186/3; 240/3

Query Match 52.1%; Score 49; DB 2; Length 531;
Best Local Similarity 61.5%; Pred. No. 8;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 WKRTPPCGSRRCG 15
:|||||
Db 335 WKCVAPMGKRRCC 347
:|||||

RESULT 4
S15146
hypothetical protein 424 - phage Pfl
C:Species: phage Pfl
C>Date: 18-Feb-1994 #sequence_revision 19-Jan-1996 #text_change 08-Oct-1999
C:Accession: S15146; S20702
R:Hill, D.P.; Short, N.J.; Perham, R.N.; Petersen, G.B.
J. Mol. Biol. 218, 349-364, 1991
A:Title: DNA sequence of the filamentous bacteriophage Pfl.
A:Reference number: S15140; MUID:91186399
A:Accession: S15146
A:Molecule type: DNA
A:Residues: 1-424 <JMO>
A:Cross-references: EMBL:X52107; NID:g14829; PIDN:CAA36334.1; PID:g579082
A:Experimental source: ATCC 25102-B1
C:Genetics:
A:Start codon: GTG

Query Match 48.9%; Score 46; DB 2; Length 424;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 DWKRTPPCGSRRCG 15
:|||||
Db 370 DWKOTVTCAGRQAG 383
:|||||

RESULT 5
T17409
Polyketide synthase type I - Streptomyces venezuelae
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:53:30 ; Search time 99.87 Seconds
(without alignments)
9,531 Million cell updates/sec

Title: us-09-236-468a-2_COPY_435_449

Perfect score: 94
Sequence: 1 VDWKRTPPGSRRCG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 182106

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	94	100.0	550	2 A57519	parathyroid hormon
2	50	53.2	446	2 PC4426	coronafacic acid s
3	49	52.1	531	2 T33099	hypothetical prote
4	46	48.9	424	2 S15146	hypothetical prote
5	44.5	47.3	4613	3 T17409	polyketide synthas
6	44	46.8	289	2 A34783	myogenesis protein
7	44	46.8	543	2 T24250	hypothetical prote
8	44	46.8	581	2 A49073	HSN motor neurons
9	43	45.7	664	1 TNBE70	70.5K alpha trans
10	42	44.7	74	2 T34608	hypothetical prote
11	42	44.7	332	2 T23653	hypothetical prote
12	42	44.7	486	2 T39456	zinc finger protei
13	42	44.7	537	2 T40247	hypothetical prote
14	42	44.7	604	2 T27870	RNA-binding / ran
15	41.5	44.1	932	2 I52527	PACE4A - mouse (fr
16	41.5	44.1	937	2 I53282	gene PACB4 protein
17	41.5	44.1	969	1 A39490	subtilisin-like pr
18	41	43.6	261	2 JQ0137	hypothetical 30.1k
19	41	43.6	287	2 C34783	myogenesis protein
20	41	43.6	372	2 A42778	agglutinin precurs
21	41	43.6	483	2 T47974	hypothetical prote
22	41	43.6	1106	2 T13938	gene shuttle craft
23	40.5	43.1	240	2 T37122	hypothetical prote
24	40.5	43.1	266	2 T10609	hypothetical prote
25	40.5	43.1	382	2 T10607	hypothetical prote
26	40.5	43.1	468	2 T10595	hypothetical prote
27	40.5	43.1	1214	2 H75034	reverse gyrase (to
28	40	42.6	103	1 W7WL5	E7 protein - human
29	40	42.6	103	1 W7WL5	E7 protein - human

30	40	42.6	126	2 T36364	hypothetical prote
31	40	42.6	139	2 T05847	hypothetical prote
32	40	42.6	398	2 A57510	mu opioid receptor
33	40	42.6	514	1 P1WL47	L1 protein - human
34	40	42.6	649	2 I64847	epithelial sodium
35	40	42.6	649	2 I38204	epithelial amilor
36	40	42.6	697	1 KIRBGC	protein kinase C (
37	40	42.6	697	1 KIRBGC	protein kinase C (
38	40	42.6	697	2 JN0548	protein kinase C (
39	40	42.6	962	2 JC5571	subtilisin-like pr
40	40	42.6	975	2 JC5570	subtilisin-like pr
41	40	42.6	1184	2 S50832	atrophin-1 - human
42	40	42.6	1234	1 NBMSH	complement factor
43	39.5	42.0	212	2 T43956	transforming trans
44	39.5	42.0	285	2 E75507	hypothetical prote
45	39.5	42.0	302	1 WMBE1E	secretory compleme

ALIGNMENTS

RESULT 1

A57519
parathyroid hormone receptor 2 precursor - human
N:Alternate names: PTHR2 receptor
C:Species: Homo sapiens (man)
C:Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 24-Nov-1999
C:Accession: A57519
R:Usdin, T.B.; Gruber, C.; Bonner, T.I.
J. Biol. Chem. 270, 15455-15458, 1995
A:Title: Identification and functional expression of a receptor selectively recognizi
A:Reference number: A57519; MUID:95318121
A:Accession: A57519
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-550 <USD>
A:Cross-references: GB:U25128; NID:g887966; PIDN:AAC50157.1; PID:g887967
C:Genetics:
A:Gene: GDB:PTHR2; PTHR2R
A:Cross-references: GDB:731977; OMIM:601469
A:Map position: 2q33-2q33
C:Superfamily: glucagon receptor
C:Keywords: hormone receptor

Query Match 100.0%; Score 94; DB 2; Length 550;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDWKRTPPGSRRCG 15

Db 435 VDWKRTPPGSRRCG 449

RESULT 2

PC4426
coronafacic acid synthetase component cfa5 [imported] - Pseudomonas syringae (fragmen
N:Contains: acyl-CoA ligase (EC 6.2.1.-)
C:Species: Pseudomonas syringae
C:Date: 24-Jan-1998 #sequence_revision 13-Mar-1998 #text_change 20-Jun-2000
C:Accession: PC4426
R:Penfold, C.N.; Bender, C.L.; Turner, J.G.
Gene 183, 167-173, 1996
A:Title: Characterisation of genes involved in biosynthesis of coronafacic acid, the
A:Reference number: JC5745; MUID:97149295
A:Accession: PC4426
A:Molecule type: DNA
A:Residues: 1-446 <PEN>
A:Cross-references: GB:U56980
A:Note: the authors translated the codon AAC for residue 2 as Asp and AAC for residue
C:Genetics:
A:Gene: cfa5
C:Keywords: acid-Chiol ligase

THIS PAGE BLANK (USPTO)

; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Disulfide-bonds
; LOCATION: 1..9
; IDENTIFICATION METHOD: by experiment
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 18
; IDENTIFICATION METHOD: by experiment
; OTHER INFORMATION: /label= Xaa
; OTHER INFORMATION: /note= "beta-(2-thienyl)-L-alanine"
US-08-382-013A-32

Query Match 42.6%; Score 40; DB 1; Length 20;
Best Local Similarity 75.0%; Pred. No. 8.2;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 DWKRTPPC 9
||| |||
Db 2 DWKGTSPC 9

RESULT 15
US-08-476-169-32
; Sequence 32, Application US/08476169
; Patent No. 5677280
; GENERAL INFORMATION:
; APPLICANT: Barrett, Ronald W.
; APPLICANT: England, Bruce
; APPLICANT: Schatz, Peter
; APPLICANT: Sloan, Derek
; APPLICANT: Chen, Min-Jia
; TITLE OF INVENTION: Peptides and Compounds That Bind to the IL-5
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Affymax Technologies, N.V.
; STREET: 4001 Miranda Ave.
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,169
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens, Lauren L.
; REGISTRATION NUMBER: 36,691
; REFERENCE/DOCKET NUMBER: 1088.2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-496-2300
; TELEFAX: 415-424-0832
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-476-169-32

Query Match 41.5%; Score 39; DB 1; Length 15;
Best Local Similarity 66.7%; Pred. No. 8.7;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VDWKRTPPC 9
||| |||
Db 6 VDWARCPTC 14

Search completed: November 8, 2000, 08:49:21
Job time: 114 sec

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933.750C
FILING DATE: September 23, 1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LNDNOT03
CLONE: 1574624
US-08-933-750C-27

Query Match 43.6%; Score 41; DB 2; Length 174;
Best Local Similarity 58.8%; Pred. No. 44;
Matches 10; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 1 VDW--KRTPPCGSRRCG 15
||| ||| ||| |||
Db 146 VDMCFVRGPPKGRRCG 162

RESULT 13
US-08-382-013A-31
Sequence 31, Application US/08382013A
Patent No. 5631222
GENERAL INFORMATION:
APPLICANT: Shibata, Kenji
APPLICANT: Suzawa, Toshiyuki
APPLICANT: Yamasaki, Motoo
APPLICANT: Tanaka, Takeo
APPLICANT: Tsukuda, Eiichi
APPLICANT: Yamada, Koji
APPLICANT: Ohno, Tetsuji
TITLE OF INVENTION: Endothelin-Antagonizing Peptide
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/382,013A
FILING DATE: 10-FEB-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 155031/93
FILING DATE: 25-JUN-1993
ATTORNEY/AGENT INFORMATION:

NAME: Lippert, Nels T.
REGISTRATION NUMBER: 25888
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-819-8200
TELEFAX: 212-354-8113
TELEX: 233188
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 18
IDENTIFICATION METHOD: by experiment
OTHER INFORMATION: /label=Xaa
OTHER INFORMATION: /note="beta-(2-thienyl)-L-alanine"
US-08-382-013A-31

Query Match 42.6%; Score 40; DB 1; Length 20;
Best Local Similarity 75.0%; Pred. No. 8.2;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 DWKRTPPC 9
||| |||
Db 2 DWKGTSPC 9

RESULT 14
US-08-382-013A-32
Sequence 32, Application US/08382013A
Patent No. 5631222
GENERAL INFORMATION:
APPLICANT: Shibata, Kenji
APPLICANT: Suzawa, Toshiyuki
APPLICANT: Yamasaki, Motoo
APPLICANT: Tanaka, Takeo
APPLICANT: Tsukuda, Eiichi
APPLICANT: Yamada, Koji
APPLICANT: Ohno, Tetsuji
TITLE OF INVENTION: Endothelin-Antagonizing Peptide
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/382,013A
FILING DATE: 10-FEB-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 155031/93
FILING DATE: 25-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Lippert, Nels T.
REGISTRATION NUMBER: 25888
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-819-8200
TELEFAX: 212-354-8113
TELEX: 233188
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:

```

; MOLECULE TYPE: protein
US-08-284-941-2

Query Match 44.1%; Score 41.5; DB 2; Length 969;
Best Local Similarity 32.1%; Pred. No. 1.8e+02;
Matches 9; Conservative 0; Mismatches 4; Indels 15; Gaps 1;

Qy 2 DWKRTPPCGS-----RRC 14
Db 874 DWKCVPCACGEGFYPEMPGLPHKVCRC 901

RESULT 10
US-08-447-642-2
; Sequence 2, Application US/08447642
; Patent No. 5989890
; GENERAL INFORMATION:
; APPLICANT: BARR, PHILIP J
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM
; STREET: FIVE PALO ALTO SQUARE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447.642
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/284,941
; FILING DATE: 2 August 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: NEELEY PH.D., RICHARD L.
; REGISTRATION NUMBER: 30092
; REFERENCE/DOCKET NUMBER: CHIR-009/01US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 843-5070
; TELEFAX: (415) 857-0663
; TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 969 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-447-642-2

Query Match 44.1%; Score 41.5; DB 2; Length 969;
Best Local Similarity 32.1%; Pred. No. 1.8e+02;
Matches 9; Conservative 0; Mismatches 4; Indels 15; Gaps 1;

Qy 2 DWKRTPPCGS-----RRC 14
Db 874 DWKCVPCACGEGFYPEMPGLPHKVCRC 901

RESULT 11
PCT-US93-02147A-2
; Sequence 2, Application PC/TUS9302147A
; GENERAL INFORMATION:
; APPLICANT: BARR, PHILIP J

```

```

; APPLICANT: KIEFER, MICHAEL C
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM
; STREET: FIVE PALO ALTO SQUARE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/02147A
; FILING DATE: 19930309
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/848,629
; FILING DATE: 09-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: NEELEY PH.D., RICHARD L.
; REGISTRATION NUMBER: 30092
; REFERENCE/DOCKET NUMBER: CHIR-009/00US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-7622
; TELEFAX: (415) 857-0663
; TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 969 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US93-02147A-2

Query Match 44.1%; Score 41.5; DB 4; Length 969;
Best Local Similarity 32.1%; Pred. No. 1.8e+02;
Matches 9; Conservative 0; Mismatches 4; Indels 15; Gaps 1;

Qy 2 DWKRTPPCGS-----RRC 14
Db 874 DWKCVPCACGEGFYPEMPGLPHKVCRC 901

RESULT 12
US-08-933-750C-27
; Sequence 27, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,174
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/137,800
FILING DATE: 19-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/084,848
FILING DATE: 29-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-107673
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Conus striatus
US-08-487-174-22

Query Match 44.7% Score 42; DB 1; Length 15;
Best Local Similarity 63.6%; Pred. No. 3.3;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 RTPPCGSRRCG 15
| | | | |
Db 5 RNPACESHRCG 15

RESULT 8
US-08-480-750-22
Sequence 22, Application US/08480750
Patent No. 5633347
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Hillyard, David R.
APPLICANT: Macintosh, J. Michael
APPLICANT: Santos, Aneurfino S.
TITLE OF INVENTION: Conotoxin Peptides
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,750
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/137,800
FILING DATE: 19-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/084,848
FILING DATE: 29-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-107673
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Conus striatus
US-08-480-750-22

Query Match 44.7% Score 42; DB 1; Length 15;
Best Local Similarity 63.6%; Pred. No. 3.3;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 RTPPCGSRRCG 15
| | | | |
Db 5 RNPACESHRCG 15

RESULT 9
US-08-284-941-2
Sequence 2, Application US/08284941
Patent No. 5863756
GENERAL INFORMATION:
APPLICANT: BARR, PHILIP J
APPLICANT: KIEFER, MICHAEL C
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND
TITLE OF INVENTION: PACE 4.1 GENE AND POLYPEPTIDES IN CELLS
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM
STREET: FIVE PALO ALTO SQUARE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,941
FILING DATE: 2 August 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: NEELEY PH.D., RICHARD L.
REGISTRATION NUMBER: 30092
REFERENCE/DOCKET NUMBER: CHIR-009/01US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 843-5070
TELEFAX: (415) 857-0663
TELEX: 380816 COOLEY PA
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 969 amino acids
TYPE: amino acid
TOPOLOGY: linear

Best Local Similarity 56.2%; Pred. No. 2.9e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

Qy 1 VDKRTPPC-GSRRCG 15

Db 1944 IDWKRLPAEGSERTG 1959

RESULT 5

US-08-137-800-22
; Sequence 22, Application US/08137800
; Patent No. 5514774
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hillyard, David R.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Santos, Ameurfin S.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,800
; FILING DATE: 19-OCT-1993
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24260-104763
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-8300
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus striatus
; US-08-137-800-22

Query Match 44.7%; Score 42; DB 1; Length 15;
Best Local Similarity 63.6%; Pred. No. 3.3;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 RTPCGSRRCG 15

Db 5 RNPACESHRCG 15

RESULT 6

US-08-477-383-22
; Sequence 22, Application US/08477383
; Patent No. 5589340
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hillyard, David R.
; APPLICANT: Macintosh, J. Michael

; APPLICANT: Santos, Ameurfin S.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,383
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/137,800
; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/084,848
; FILING DATE: 29-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24260-107673
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus striatus
; US-08-477-383-22

Query Match 44.7%; Score 42; DB 1; Length 15;
Best Local Similarity 63.6%; Pred. No. 3.3;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 RTPCGSRRCG 15

Db 5 RNPACESHRCG 15

RESULT 7

US-08-487-174-22
; Sequence 22, Application US/08487174
; Patent No. 5595972
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hillyard, David R.
; APPLICANT: Macintosh, J. Michael
; APPLICANT: Santos, Ameurfin S.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005

Db 435 VDWKRTPPCGSRRCG 449

RESULT 2

PCT-US95-07085-2
; Sequence 2, Application PC/TUS9507085
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone Receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ

; COUNTRY: USA
; ZIP: 07068-1739
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07085
; FILING DATE: 05-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-393
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-07085-2

Query Match 100.0%; Score 94; DB 4; Length 541;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDWKRTPPCGSRRCG 15
|||||

Db 435 VDWKRTPPCGSRRCG 449

RESULT 3

US-08-468-011A-23
; Sequence 23, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone Receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland

; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 59 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-011A-23

Query Match 73.4%; Score 69; DB 3; Length 59;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDWKRTPPCGS 11
|||||

Db 49 VDWKRTPPCGS 59

RESULT 4

US-09-320-878-1
; Sequence 1, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4551
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
; US-09-320-878-1

Query Match

47.3%; Score 44.5; DB 3; Length 4551;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:49:19 ; Search time 97.15 Seconds
(without alignments)
2.588 Million cell updates/sec

Title: US-09-236-468A-2_COPY_435_449

Perfect score: 94
Sequence: 1 VDKRTPCGRRCG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 164575

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	100.0	541	3	US-08-468-011A-2
2	94	100.0	541	4	PCT-US95-07085-2
3	69	73.4	59	3	US-08-468-011A-23
4	44.5	47.3	4551	3	US-09-320-878-1
5	42	44.7	15	1	US-08-137-800-22
6	42	44.7	15	1	US-08-477-383-22
7	42	44.7	15	1	US-08-487-174-22
8	42	44.7	15	1	US-08-480-750-22
9	41.5	44.1	969	2	US-08-284-941-2
10	41.5	44.1	969	2	US-08-447-642-2
11	41.5	44.1	969	4	PCT-US93-02147A-2
12	41	43.6	174	2	US-08-933-750C-27
13	40	42.6	20	1	US-08-382-013A-31
14	40	42.6	20	1	US-08-382-013A-32
15	39	41.5	15	1	US-08-476-169-32
16	39	41.5	15	1	US-08-476-169-37
17	39	41.5	15	1	US-08-476-169-41
18	39	41.5	15	1	US-08-476-169-48
19	39	41.5	15	1	US-08-476-169-52
20	39	41.5	15	1	US-08-484-083-32
21	39	41.5	15	1	US-08-484-083-37
22	39	41.5	15	1	US-08-484-083-41
23	39	41.5	15	1	US-08-484-083-48
24	39	41.5	15	1	US-08-484-083-52
25	39	41.5	38	1	US-08-176-500-57
26	39	41.5	38	1	US-08-471-052A-57
27	39	41.5	38	1	US-08-189-331-57
28	39	41.5	38	2	US-08-471-939-57

Sequence 57, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 4, Appl
Sequence 5, Appl
Sequence 8, Appl
Sequence 5, Appl
Sequence 2, Appl
Sequence 38, Appl
Sequence 2, Appl
Sequence 46, Appl
Sequence 47, Appl
Sequence 49, Appl

ALIGNMENTS

RESULT 1
US-08-468-011A-2
; Sequence 2, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESS: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468.011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-011A-2

Query Match 100.0%; Score 94; DB 3; Length 541;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VDKRTPCGRRCG 15
|||||

A-lineage conotoxin peptide ST-1.

Polymerase chain reaction: PCR: primer: amplif; conotoxin; Conus;
inhibitor; synaptic transmission; neuromuscular junction; sodium channel;
nicotinic acetylcholine receptor; potassium channel; muscle relaxant;
myasthenia gravis; small cell lung cancer; therapy.

Conus striatus.

US5589340-A.

31-DEC-1996.

29-JUN-1993; 93US-0084848.

07-JUN-1995; 95US-0477383.

29-JUN-1993; 93US-0084848.

19-OCT-1993; 93US-0137800.

(UTAH) UNIV UTAH RES FOUND.

Cruz LJ, Hillyard DR, McIntosh JM, Olivera BM, Santos AD;

WPI; 1997-076840/07.

Identifying nucleic acid encoding A-lineage conotoxin peptide(s) by
amplification - uses primers corresponding to conserved regions in
the signal sequence and 3'-untranslated regions, useful e.g. in
treatment of small cell lung cancer

Disclosure; Column 5; 36pp; English.

W12726-W12769 represent conotoxin peptides. This sequence represents the
A-lineage conotoxin ST-1 peptide isolated from Conus striatus. These
sequences are identified using the method of the invention. The method
of the invention is for identifying DNA encoding A-lineage conotoxin
peptides by subjecting Conus nucleic acid to amplification with primer
sequences (see T59714 and T59715). The primers are specific for the
signal sequence and 3'-untranslated (3'UTR) regions of the conotoxin
gene, which are highly homologous between conotoxins, and are therefore
suitable sites for detection. A-lineage conotoxins include alpha-
conotoxins, and kappa-conotoxins. Alpha-conotoxins are powerful
inhibitors of synaptic transmission at the neuromuscular junction, and
are usually nicotinic acetylcholine receptor blockers. Kappa-conotoxins
act on the voltage sensitive sodium and potassium channels. The
conotoxins identified can be used as muscle relaxants, in the diagnosis
of myasthenia gravis, and for the treatment or diagnosis of small cell
lung cancer. For the treatment of small cell lung cancer, the conotoxin
peptides act by binding to the nicotinic receptors, and thereby blocking
the nicotine/cytosine stimulated release of the mitogen
5-hydroxytryptamine.

Sequence 15 AA;

Query Match 44.7%; Score 42; DB 18; Length 15;
Best Local Similarity 63.6%; Pred. No. 7.2;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 RTPPCGSRRCG 15
I I I I I I I I
Db 5 rnpaceshrcg 15

RESULT 15

Y57957
ID Y57957 standard; protein; 461 AA.
XX
XX AC Y57957;
XX
XX DT 24-MAR-2000 (first entry)
XX

Fukuyama-type congenital muscular dystrophy-causing protein.

XX
KW
KW
XX
OS
XX
PN
XX
PD
XX
PF
XX
PR
XX
PA
XX
DR
DR
XX
PT
PT
XX
PS
SQ

Fukuyama-type congenital muscular dystrophy-causing protein; FCMD;
detection; muscular dystrophy; diagnosis.

Homo sapiens.

JP11313682-A.

16-NOV-1999.

30-APR-1998; 98JP-0137703.

30-APR-1998; 98JP-0137703.

(SAKA) OTSUKA PHARM CO LTD.

WPI; 2000-090363/08.

N-PSDB; 256799, Z57101.

A Fukuyama-type congenital muscular dystrophy-causing protein - for
preparing its specific antibody

Claim 1; Page 15-16; 32pp; Japanese.

The present sequence represents a Fukuyama-type congenital muscular
dystrophy (FCMD)-causing protein isolated from human. Also described
in the present invention is a method for the detection of gene
abnormality for FCMD diagnosis by detecting the presence of a mutated
FCMD-causing DNA having a mutation causing functional insufficiency of
the FCMD-causing protein coded in the base sequence of 7389 nucleotides
in the gene of a person to be tested. The FCMD-causing protein is
useful in the preparation of its specific antibody.

Sequence 461 AA;

Query Match 44.7%; Score 42; DB 21; Length 461;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DMKRTTP 8
I I I I I I I I
Db 435 dwkrspp 441

Search completed: November 8, 2000, 08:51:48
Job time: 260 sec

Db 2006 idwrlpaaegsertg 2021

RESULT 10

Y77180

ID Y77180 standard; Protein; 12199 AA.

XX

AC Y77180;

XX

XX 05-JUN-2000 (first entry)

XX

XX S. venezuelae pik gene cluster-encoded protein, SEQ ID NO:6.

XX

XX Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;

KW neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase;

KW biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma,

KW chronic obstructive pulmonary disease; respiratory inflammation;

KW hypercholesterolaemia; crop protection agent.

XX

OS Streptomyces venezuelae ATCC15439.

XX

XX WO200000620-A2.

XX

XX 06-JAN-2000.

XX

XX 25-JUN-1999; 99WO-US14398.

XX

XX 26-JUN-1998; 98US-0105537.

XX

XX (MINU) UNIV MINNESOTA.

XX

XX Sherman DH, Liu H, Xue Y, Zhao L;

XX

XX WPI; 2000-160679/14.

DR

DR N-PSDB; 287285.

XX

XX Desosamine and macrolide biosynthetic gene clusters, useful for, e.g.

PT synthesis of methymycin and pikromycin -

XX

PS Claim 19; Page 315-353; 438pp; English.

XX

XX The invention relates to an isolated and purified nucleic acid segment

CC comprising a desosamine biosynthetic gene cluster, a fragment or its

CC biologically active variant, where the nucleic acid sequence is not

CC derived from the ercC gene cluster of *Saccharopolyspora erythraea* or

CC Streptomyces antibioticus. The invention also relates to a macrolide

CC biosynthetic gene cluster, or fragments thereof. The macrolide

CC biosynthetic gene cluster encodes proteins which synthesise methymycin,

CC pikromycin, neomethymycin, narbomycin or a combination of these

CC compounds. Recombinant or augmented cells comprising the desosamine

CC and/or macrolide biosynthetic gene clusters are useful for the production

CC of biologically active macrolides. The macrolide biosynthetic proteins

CC are useful for synthesis of methymycin, pikromycin, neomethymycin and

CC narbomycin. The alternative termination of polyketide synthesis may be

CC useful to prepare novel antibiotics and polyhydroxyalkanoate (PHA)

CC monomers. The compounds produced by the recombinant host cells are useful

CC as biopolymers, e.g., in packaging or biomedical applications, to

CC engineer PHA monomer synthases or to prepare biologically active agents,

CC such as chemotherapeutics, immunosuppressants, agents to treat asthma,

CC chronic obstructive pulmonary disease as well as other diseases involving

CC respiratory inflammation, cholesterol-lowering agents or macrolide-based

CC antibiotics which are active against a variety of organisms, e.g.,

CC bacteria, including multi-drug resistant pneumococci and other

CC respiratory pathogens, as well as viral parasitic pathogens, or as crop

CC protection agents (e.g., fungicides or insecticides) via expression of

CC polyketides in plants. The present sequence represents a protein

CC encoded by the macrolide biosynthetic gene cluster (pik) from

CC Streptomyces venezuelae ATCC 15439.

XX

XX Sequence 12199 AA;

XX

Query Match 47.3%; Score 44.5; DB 21; Length 12199;

Best Local Similarity 58.3%; Pred. No. 17;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 WKRTPPCGSRRC 14

Db 13 wrtctpydspac 24

Best Local Similarity 56.2%; Pred. No. 1e+03;

Matches 9; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 VDWKRTPPC-GSRRCG 15

Db 2664 idwrlpaaegsertg 2679

RESULT 11

Y76578

ID Y76578 standard; Protein; 86 AA.

XX

AC Y76578;

XX

XX 10-APR-2000 (first entry)

XX

XX Human ovarian tumor EST fragment encoded protein 74.

DE

XX

KW Expressed sequence tag; EST; human; ovarian tumor; anticancer;

KW gene therapy; treatment.

XX

XX Homo sapiens.

XX

XX DEL9817557-A1.

XX

XX 21-OCT-1999.

XX

XX 09-APR-1998; 98DE-1017557.

XX

XX 09-APR-1998; 98DE-1017557.

XX

XX (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX

XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pillarsky C, Dahl E;

XX

XX WPI; 1999-591920/51.

DR

DR N-PSDB; 277482.

XX

XX New nucleic acid sequences expressed in ovarian, and some other, cancer

PT tissues, and derived polypeptides, for treatment of ovarian cancer and

PT identification of therapeutic agents -

XX

PS Claim 25; Page 274; 310pp; German.

XX

XX This invention describes novel nucleic acid (cDNA) sequences (A) which

CC have anticancer activity and are highly expressed in ovarian tumor

CC tissue (and some also in testis and breast cancer tissue). The products

CC of the invention can be used for gene therapy. (A) are used (i) for

CC recombinant expression of polypeptides (B) and (ii) to isolate complete

CC genes. (B) are used (i) to identify agents suitable for treatment of

CC ovarian cancer; (ii) directly for treating this form of cancer (including

CC expression from gene therapy vectors) and (iii) for generation of

CC specific antibodies. (A) are identified by assembling ESTs (expressed

CC sequence tags) from a particular tissue type before comparison of

CC expression patterns. This allows a significantly longer fragment of the

CC gene to be revealed, so should reduce the number of failures associated

CC with the fact that ESTs from different libraries may represent different

CC parts of the same unknown gene, distorting the estimated frequency of

CC occurrence in a particular tissue. Y76505-Y76638 represent protein

CC fragments encoded by the human ovarian tumor cDNA library derived EST

CC fragments represented in Z77450-Z77572.

XX

XX Sequence 86 AA;

XX

Query Match 46.8%; Score 44; DB 20; Length 86;

Best Local Similarity 58.3%; Pred. No. 17;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 WKRTPPCGSRRC 14

Db 13 wrtctpydspac 24

KW chronic obstructive pulmonary disease; respiratory inflammation;
 KW hypercholesterolaemia; crop protection agent.
 OS Streptomyces venezuelae ATCC15439.

FH key Location/Qualifiers
 FT Misc-difference 1386 /note= "Encoded by GNN"
 FT Misc-difference 4581 /note= "Encoded by GNN"
 XX

PN WO200000620-A2.

XX

PD 06-JAN-2000.

XX

PF 25-JUN-1999; 99WO-US14398.

XX

PR 26-JUN-1998; 98US-0105537.

XX

PA (MINU) UNIV MINNESOTA.

XX

PI Sherman DH, Liu H, Xue Y, Zhao L;

XX

DR WPI; 2000-160679/14.

XX

DR N-PSDB; 287297.

XX

PT Desosamine and macrolide biosynthetic gene clusters, useful for, e .g.
 PT synthesis of methymycin and pikromycin -
 XX

PS Claim 19; Page 383-398; 438pp; English.

XX

CC The invention relates to an isolated and purified nucleic acid segment
 CC comprising a desosamine biosynthetic gene cluster, a fragment or its
 CC biologically active variant, where the nucleic acid sequence is not
 CC derived from the eryc gene cluster of Saccharopolyspora erythraea or
 CC Streptomyces antibioticus. The invention also relates to a macrolide
 CC biosynthetic gene cluster, or fragments thereof. The macrolide
 CC biosynthetic gene cluster encodes proteins which synthesize methymycin,
 CC pikromycin, neomethymycin, narbomycin or a combination of these
 CC compounds. Recombinant or augmented cells comprising the desosamine
 CC and/or macrolide biosynthetic gene clusters are useful for the production
 CC of biologically active macrolides. The macrolide biosynthetic proteins
 CC are useful for synthesis of methymycin, pikromycin, neomethymycin and
 CC narbomycin. The alternative termination of polyketide synthesis may be
 CC useful to prepare novel antibiotics and polyhydroxyalkanoate (PHA)
 CC monomers. The compounds produced by the recombinant host cells are useful
 CC as biopolymers, e.g., in packaging or biomedical applications, to
 CC engineer PHA monomer synthases or to prepare biologically active agents,
 CC such as chemotherapeutics, immunosuppressants, agents to treat asthma,
 CC chronic obstructive pulmonary disease as well as other diseases involving
 CC respiratory inflammation, cholesterol-lowering agents or macrolide-based
 CC antibiotics which are active against a variety of organisms, e.g.,
 CC bacteria, including multi-drug resistant pneumococci and other
 CC respiratory pathogens, as well as viral parasitic pathogens, or as crop
 CC protection agents (e.g., fungicides or insecticides) via expression of
 CC polyketides in plants. Sequences Y77190-Y77197 represent macrolide
 CC biosynthetic enzymes from Streptomyces venezuelae ATCC 15439, which
 CC are encoded by sequences 287295-287302.

XX

SQ Sequence 4613 AA;

Query Match 47.3%; Score 44.5; DB 21; Length 4613;
 Best Local Similarity 56.2%; Pred. No. 4.4e+02;
 Matches 9; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

Oy 1 VDWKRTPPC-GSRRCG 15

Db 2006 idwkrtpaegsertg 2021

RESULT 9

Y77200

Y77200 standard; Protein; 4613 AA.

XX

AC Y77200;

XX

DT 05-JUN-2000 (first entry)

XX

DE S. venezuelae pik (macrolide biosynthesis) gene cluster protein #1.
 KW Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;
 KW neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase;
 KW biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma,
 KW chronic obstructive pulmonary disease; respiratory inflammation;
 KW hypercholesterolaemia; crop protection agent.

XX

OS Streptomyces venezuelae ATCC15439.

XX

PN WO200000620-A2.

XX

PD 06-JAN-2000.

XX

PF 25-JUN-1999; 99WO-US14398.

XX

PR 26-JUN-1998; 98US-0105537.

XX

PA (MINU) UNIV MINNESOTA.

XX

PI Sherman DH, Liu H, Xue Y, Zhao L;

XX

DR WPI; 2000-160679/14.

XX

DR N-PSDB; 287318.

XX

PT Desosamine and macrolide biosynthetic gene clusters, useful for, e .g.
 PT synthesis of methymycin and pikromycin -
 XX

PS Disclosure; Figure 31; 438pp; English.

XX

CC The invention relates to an isolated and purified nucleic acid segment
 CC comprising a desosamine biosynthetic gene cluster, a fragment or its
 CC biologically active variant, where the nucleic acid sequence is not
 CC derived from the eryc gene cluster of Saccharopolyspora erythraea or
 CC Streptomyces antibioticus. The invention also relates to a macrolide
 CC biosynthetic gene cluster, or fragments thereof. The macrolide
 CC biosynthetic gene cluster encodes proteins which synthesize methymycin,
 CC pikromycin, neomethymycin, narbomycin or a combination of these
 CC compounds. Recombinant or augmented cells comprising the desosamine
 CC and/or macrolide biosynthetic gene clusters are useful for the production
 CC of biologically active macrolides. The macrolide biosynthetic proteins
 CC are useful for synthesis of methymycin, pikromycin, neomethymycin and
 CC narbomycin. The alternative termination of polyketide synthesis may be
 CC useful to prepare novel antibiotics and polyhydroxyalkanoate (PHA)
 CC monomers. The compounds produced by the recombinant host cells are useful
 CC as biopolymers, e.g., in packaging or biomedical applications, to
 CC engineer PHA monomer synthases or to prepare biologically active agents,
 CC such as chemotherapeutics, immunosuppressants, agents to treat asthma,
 CC chronic obstructive pulmonary disease as well as other diseases involving
 CC respiratory inflammation, cholesterol-lowering agents or macrolide-based
 CC antibiotics which are active against a variety of organisms, e.g.,
 CC bacteria, including multi-drug resistant pneumococci and other
 CC respiratory pathogens, as well as viral parasitic pathogens, or as crop
 CC protection agents (e.g., fungicides or insecticides) via expression of
 CC polyketides in plants. Sequences Y77200-Y77203 and Y80997 represent
 CC proteins encoded by the macrolide biosynthetic gene cluster (pik) from
 CC Streptomyces venezuelae ATCC 15439, as given in figure 31.

XX

SQ Sequence 4613 AA;

Query Match 47.3%; Score 44.5; DB 21; Length 4613;
 Best Local Similarity 56.2%; Pred. No. 4.4e+02;
 Matches 9; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

Oy 1 VDWKRTPPC-GSRRCG 15

:||||| | | | |

ID R39870 standard; peptide; 28 AA.
 AC R39870;
 XX
 DT 19-JAN-1994 (first entry)
 XX
 DE C peptide RV-C3, residues 52-78.
 XX
 XX Rubella virus; RV; C protein; antibody; mammal; vaccine; rubella;
 KW neutralising; cell mediated; capsid; immune response; mumps; measles.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 28
 FT /note= "May be absent"
 XX
 XX WO9314206-A.
 PD 22-JUL-1993.
 XX
 XX 20-JAN-1993; 93WO-CA00014.
 XX
 XX 20-JAN-1992; 92GB-0001139.
 XX
 PA (CONN-) CONNAUGHT LAB LTD.
 XX
 PI Chong P, Gillam S, Tingle A;
 XX
 DR WPI; 1993-243221/30.
 XX
 XX Synthetic peptide(s) having at least one antigenic determinant of
 PT a rubella virus protein - useful for producing vaccine, and also
 PT to detect associated antibodies to treat associated auto-immune
 PT disorders, etc.
 XX
 PS Table 3; Page 44; 68pp; English.
 XX
 XX The sequences given in R39868-78 represent rubella virus (RV) capsid
 CC (C) protein peptide fragments. These peptides are capable of eliciting
 CC high titres of antibodies against RV in mammals. They may be used
 CC in vaccines to elicit neutralising antibodies and a cell mediated
 CC immune response against RV. They may be used as one component of a
 CC multivalent vaccine, pref. one providing protection against rubella,
 CC mumps and measles.
 XX
 SQ Sequence 28 AA;

Query Match 47.9%; Score 45; DB 14; Length 28;
 Best Local Similarity 75.0%; Pred. No. 4.7;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 DWKRTTPC 9
 || || ||
 Db 21 dwsrappc 28

RESULT 7

ID Y67201 standard; protein; 4551 AA.
 AC Y67201;
 XX
 XX Y67201;
 XX
 DT 23-MAR-2000 (first entry)
 XX

DE Narbonolide synthase subunit 1 (PICAI) protein sequence.
 XX

KW Narbonolide polyketide synthase; PKS; narbonolide synthase subunit 1;
 KW PICAI; antibiotic production; narbomycin; picromycin; ketolide.
 XX
 OS Streptomyces venezuelae.
 XX

PN WO9961599-A2.
 XX
 PD 02-DEC-1999.
 XX
 PF 19-MAY-1999; 99WO-US11814.
 XX
 XX 27-MAY-1999; 99WO-US11814.
 XX
 PR 28-MAY-1998; 98US-0087080.
 PR 28-AUG-1998; 98US-0141908.
 PR 22-SEP-1998; 98US-0100880.
 PR 08-FEB-1999; 99US-0119139.
 XX
 XX (KOSA-) KOSAN BIOSCIENCES INC.
 PA
 XX
 XX Ashley G, Betlach MC, Betlach M, McDaniel R, Tang L;
 PI
 XX
 XX WPI; 2000-072618/06.
 DR
 DR N-PSDB; Z56001.
 XX
 PT New recombinant DNA encoding a domain of narbonolide polyketide
 PT synthase, for production of ketolide antibiotics -
 XX
 XX Example 2; Page 11-12; 98pp; English.
 XX
 CC This is the Streptomyces venezuelae narbonolide synthase subunit 1,
 CC PICAI protein sequence. The invention relates to recombinant DNA
 CC containing a coding sequence for a narbonolide polyketide synthase (PKS).
 CC Polyketides are compounds synthesised from 2-carbon units through a
 CC series of condensations and subsequent modifications. Modular PKSs are
 CC responsible for the production of many antibiotics including picromycin.
 CC The narbonolide PKS consists of a loading module, six extender modules,
 CC and two thioester domains. Four proteins make up the narbonolide PKS
 CC (PICAI, PICAI1, PICAI2 and PICAI3). PICAI includes the loading module
 CC and extender modules 1 and 2, PICAI1 includes extender modules 3 and 4,
 CC PICAI2 includes extender module 5 and PICAI3 includes extender module 6
 CC and a type II thioesterase domain. The second type II thioesterase
 CC domain is found on the PICB protein. The nucleotide sequences encoding
 CC all of these proteins can be isolated in recombinant form from the
 CC recombinant cosmid pKOS023-27 (see Z56001). Narbonolide is
 CC desosaminylated in S. venezuelae to yield narbomycin, and the desosamine
 CC transferase enzyme is required for this conversion, and the desosamine
 CC biosynthetic genes are also found in cosmid pKOS023-27. The recombinant
 CC DNA of the invention is used to express, in transformed cells,
 CC narbonolide (or its derivatives) or other ketolides (particularly
 CC hybrids), which may then be converted (e.g. by other enzymes
 CC recombinantly expressed in the same hosts) to polyketide antibiotics or
 CC their intermediates. The antibiotics are useful in human or veterinary
 CC medicine.
 XX
 SQ Sequence 4551 AA;

Query Match 47.3%; Score 44.5; DB 21; Length 4551;
 Best Local Similarity 56.2%; Pred. No. 4.3e+02;
 Matches 9; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 VDMKRTTPC-GSRRCG 15
 :||| | ||| |
 Db 1944 idwrlpaaegsertg 1959

RESULT 8

ID Y77192 standard; protein; 4613 AA.
 XX
 AC Y77192;
 XX
 XX Y77192;
 XX
 DT 05-JUN-2000 (first entry)
 XX

DE S. venezuelae macrolide biosynthetic enzyme PikAI, SEQ ID NO:31.
 XX

KW Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;
 KW neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase;
 KW biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma,
 XX

```
ID Y54146 standard; Protein; 386 AA.
XX
AC Y54146;
XX
DT 27-MAR-2000 (first entry)
XX
DE Amino acid sequence of the motilin receptor splice variant MTL-R1B.
XX
KW Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
KW spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
KW functional defect; neurological disorder; scleroderma; colonoscopy;
KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;
KW infection; stress-related motility disorder; psychogenic disorder;
KW gastroparesis; gastro-oesophageal reflux disease; constipation;
KW chronic idiopathic pseudo obstruction; acute faecal impaction;
KW postoperative ileus; gallstones; infantile colic; diarrhoea;
KW irritable bowel syndrome; non-ulcer dyspepsion; non-cardiac chest pain;
KW endoscopy; duodenal intubation.
XX
OS Homo sapiens.
XX
PN WO9964436-A1.
XX
PD 16-DEC-1999.
XX
PF 08-JUN-1999; 99WO-US12773.
XX
PR 12-JUN-1998; 98US-0089098.
XX
PA (MERI ) MERCK & CO INC.
XX
PI Feighner SD, Patchett AA, Tan C, McKee K, MacNeill D, Howard AD;
PI Pong S, Smith RG;
XX
DR WPI; 2000-105868/09.
DR N-PSDB; Z45404.
XX
XX Novel receptor protein for screening compounds used in treating
PT irritable bowel syndrome, constipation and other gastric conditions -
XX
PS Claim 5; Fig 5; 44pp; English.
XX
XX The present sequence represents splice variant MTL-R1B of the motilin
CC receptor. The gene encodes a G-protein coupled receptor, and is
CC designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,
CC MTL-R1A (see Y54145) and MTL-R1B (see Y54146). MTL-R1A is a functional
CC seven transmembrane domain form, and MTL-R1B is a truncated five
CC transmembrane domain. The MTL-R1 proteins are used to identify agonists
CC and antagonists which can be used for treating gastric motility
CC disorders, functional defects, disorders secondary to neurological
CC disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
CC dysmotility, diabetes, infections, stress-related motility disorders,
CC psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,
CC constipation, chronic idiopathic pseudo obstruction, acute faecal
CC impaction, postoperative ileus, gallstones, infantile colic, irritable
CC bowel syndrome, non-ulcer dyspepsion, non-cardiac chest pain and
CC diarrhoea. They can also be used in the preparation for colonoscopy,
CC endoscopy and duodenal intubation.
XX
SQ Sequence 386 AA;

Query Match 52.1%; Score 49; DB 21; Length 386;
Best Local Similarity 56.3%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 WKRTPPCGSRRRC 14
| | | | |
Db 19 wpa1ppcderrc 30

RESULT 5
Y54145
```

```
ID Y54145 standard; Protein; 412 AA.
XX
AC Y54145;
XX
DT 27-MAR-2000 (first entry)
XX
DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.
XX
KW Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
KW spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
KW functional defect; neurological disorder; scleroderma; colonoscopy;
KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;
KW infection; stress-related motility disorder; psychogenic disorder;
KW gastroparesis; gastro-oesophageal reflux disease; constipation;
KW chronic idiopathic pseudo obstruction; acute faecal impaction;
KW postoperative ileus; gallstones; infantile colic; diarrhoea;
KW irritable bowel syndrome; non-ulcer dyspepsion; non-cardiac chest pain;
KW endoscopy; duodenal intubation.
XX
OS Homo sapiens.
XX
PN WO9964436-A1.
XX
PD 16-DEC-1999.
XX
PF 08-JUN-1999; 99WO-US12773.
XX
PR 12-JUN-1998; 98US-0089098.
XX
PA (MERI ) MERCK & CO INC.
XX
PI Feighner SD, Patchett AA, Tan C, McKee K, MacNeill D, Howard AD;
PI Pong S, Smith RG;
XX
DR WPI; 2000-105868/09.
DR N-PSDB; Z45403.
XX
XX Novel receptor protein for screening compounds used in treating
PT irritable bowel syndrome, constipation and other gastric conditions -
XX
PS Claim 3; Fig 3; 44pp; English.
XX
XX The present sequence represents splice variant MTL-R1A of the motilin
CC receptor. The gene encodes a G-protein coupled receptor, and is
CC designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,
CC MTL-R1A (see Y54145) and MTL-R1B (see Y54146). MTL-R1A is a functional
CC seven transmembrane domain form, and MTL-R1B is a truncated five
CC transmembrane domain. The MTL-R1 proteins are used to identify agonists
CC and antagonists which can be used for treating gastric motility
CC disorders, functional defects, disorders secondary to neurological
CC disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
CC dysmotility, diabetes, infections, stress-related motility disorders,
CC psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,
CC constipation, chronic idiopathic pseudo obstruction, acute faecal
CC impaction, postoperative ileus, gallstones, infantile colic, irritable
CC bowel syndrome, non-ulcer dyspepsion, non-cardiac chest pain and
CC diarrhoea. They can also be used in the preparation for colonoscopy,
CC endoscopy and duodenal intubation.
XX
SQ Sequence 412 AA;

Query Match 52.1%; Score 49; DB 21; Length 412;
Best Local Similarity 58.3%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 WKRTPPCGSRRRC 14
| | | | |
Db 19 wpa1ppcderrc 30

RESULT 6
R39870
```


PT hyper-calcaemia, hypo- or hyper-phosphatemia, kidney stones, etc
PS Claim 9; Fig 1A-E; 62pp; English.
XX
XX A novel 7-transmembrane receptor (W12695) has been identified as a
CC human G-protein parathyroid hormone (PTH) receptor, designated
CC HTDG74. It shows 48.2% homology to the human PTH receptor. Its
CC amino acid sequence was deduced from a cDNA clone (759619) isolated
CC from a human T cell lymphoma tissue cDNA library. Recombinant
CC HTDG74 can be produced in transformed host cells and used to
CC screen for (ant)agonist cpds. Agonists may be used to prevent or
CC treat e.g. hypocalcaemia, hyperphosphataemia, hypoparathyroidism
CC and chronic tetany by stimulating an increase in serum calcium
CC levels. Antagonists can be used to inhibit the receptor e.g. for
CC the treatment of osteoporosis, hypercalcaemia, hypoparathyroidism,
CC hypophosphataemia, kidney stone, nephroliasis.
XX
SQ Sequence 541 AA;

Query Match 100.0%; Score 94; DB 18; Length 541;
Best Local Similarity 100.0%; Pred. No. 9.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDWKRTPPCGSRRCG 15
Db 435 vdwkrtppcgssrrcg 449
|||||

RESULT 2
Y29131
ID Y29131 standard; Protein; 174 AA.
XX
AC Y29131;
XX
DT 25-OCT-1999 (first entry)
XX
DE Amino acid sequence of a virulence factor encoded by ORF7584.
XX
KW Human pathogen; virulence polypeptide; virulence factor;
KW pathogenic infection; Pseudomonas aeruginosa infection.
XX
OS Pseudomonas aeruginosa.
XX
PN WO9927129-A1.
XX
PD 03-JUN-1999.
XX
PF 25-NOV-1998; 98WO-US25247.
XX
PR 25-NOV-1997; 97US-0066517.
XX
PS (GEO) GEN HOSPITAL CORP.
XX
PA Ausubel F, Cao H, Drenkard E, Goodman HM, Mahajan-Wiklos S;
PI Rahme LG, Tan M, Tsongalis J;
XX
DR WPI; 1999-357851/30.
XX
PT Virulence factors useful in developing disease treatments
XX
PS Disclosure; Fig 4; 228pp; English.
XX
XX The present sequence represents a Pseudomonas aeruginosa polypeptide
CC sequence. P. aeruginosa is an opportunistic human pathogen present in
CC soil water and plants. The specification describes virulence polypeptides
CC and nucleic acid sequence encoding such polypeptides. These sequences
CC can be used to identify a compound which is capable of decreasing the
CC expression of a pathogenic virulence factor. Compounds that inhibit
CC the expression or activity of virulence factor polypeptides can be
CC used to treat pathogenic infections, especially where the infection
CC is a P. aeruginosa infection.
CC
CC note: the sequences given in the specification were poorly legible, and

CC In some instances assumptions were made as to the identity of the
CC residue; it is therefore possible that the sequence given below is
CC not entirely correct.
XX
SQ Sequence 174 AA;

Query Match 54.3%; Score 51; DB 20; Length 174;
Best Local Similarity 61.5%; Pred. No. 3.3;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 WKRTPPCGSRRCG 15
Db 44 wpatppcgssag 56
|||||

RESULT 3
R75282
ID R75282 standard; peptide; 15 AA.
XX
AC R75282;
XX
DT 21-DEC-1995 (first entry)
XX
DE A-lineage conotoxin ST-1 peptide.
XX
KW Alpha conotoxin; neuromuscular; synapse; signal transmission.
XX
OS Conus striatus.
XX
PN WO9511256-A1.
XX
PD 27-APR-1995.
XX
PE 19-OCT-1994; 94WO-US11927.
XX
PR 19-OCT-1993; 93US-0137800.
XX
PA (UTAH) UNIV UTAH RES FOUND.
XX
PI Cruz LJ, Hillyard DR, McIntosh JM, Olivera BM, Santos AD;
XX
DR WPI; 1995-170189/22.
XX
PT New A-lineage conotoxin peptide(s) - which inhibit synaptic
PT transmission at the neuromuscular junction or are active against
PT potassium or sodium channels
XX
PS Claim 1; Page 45; 66pp; English.
XX
XX The kappa-conotoxin, alpha conotoxin and alpha-like conotoxin
CC peptides all belong to a group of peptides known as the A-lineage
CC conotoxin peptides. The A lineage conotoxin peptides have a wide
CC variety of pharmacological uses. The A-lineage conotoxin peptides claimed
CC (R75264-R75293) are useful for the inhibition of synaptic transmission
CC at neuromuscular junctions by blocking nicotinic acetyl choline
CC receptors and they also have activity against voltage-gated Na and K
CC channels.
XX
SQ Sequence 15 AA;

Query Match 53.2%; Score 50; DB 16; Length 15;
Best Local Similarity 72.7%; Pred. No. 0.56;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 RTPPCGSRRCG 15
Db 5 rnpacgshrcg 15
|||||

RESULT 4
Y54146

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:51:46 : Search time 138.73 Seconds
(without alignments)
3.697 Million cell updates/sec

Title: US-09-236-468A-2_COPY_435_449
Perfect score: 94
Sequence: 1 VDMKRTPPCGSRRCG 15

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36.*
1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94	100.0	541	18 W12695	G-protein parathyr
2	51	54.3	174	20 Y29131	Amino acid sequenc
3	50	53.2	15	16 R75282	A-lineage conotoxi
4	49	52.1	386	21 Y54146	Amino acid sequenc
5	49	52.1	412	21 Y54145	Amino acid sequenc
6	45	47.9	28	14 R39870	C peptide RV-C3, r
7	44.5	47.3	4551	21 Y67201	Narbornolide syntha
8	44.5	47.3	4613	21 Y77192	S. venezuelae macr
9	44.5	47.3	4613	21 Y77200	S. venezuelae pik
10	44.5	47.3	12199	21 Y77180	S. venezuelae pik
11	44	46.8	86	20 Y76578	Human ovarian tumo
12	43	45.7	232	21 Y50923	Human fetal brain

13	42	44.7	15	18 W24889	Predatory cone sna
14	42	44.7	15	18 W12744	A-lineage conotoxi
15	42	44.7	461	21 Y57957	Fukuyama-type cong
16	41.5	44.1	969	14 R41662	Paired basic amino
17	41	43.6	125	21 Y92245	Murine zrnpl (ribo
18	41	43.6	158	21 Y92244	Human zrnpl (ribo
19	40	42.6	390	21 Y68885	A murine mu-opioid
20	40	42.6	391	21 Y68886	A murine mu-opioid
21	40	42.6	392	21 Y68879	A murine mu-opioid
22	40	42.6	398	19 W44937	Mouse mu opiate re
23	40	42.6	398	21 Y68889	A murine mu-opioid
24	40	42.6	401	21 Y68880	A murine mu-opioid
25	40	42.6	409	21 Y68887	A murine mu-opioid
26	40	42.6	438	21 Y68877	A murine mu-opioid
27	40	42.6	444	21 Y68888	A murine mu-opioid
28	40	42.6	547	19 W69481	Rat matrix metallo
29	40	42.6	604	18 W10640	Membrane type matr
30	40	42.6	607	19 W69480	Human matrix metal
31	40	42.6	697	9 P82018	Protein kinase C-I
32	40	42.6	710	20 W80995	Human guanine nucl
33	40	42.6	1503	19 W48845	Human receptor tyr
34	39.5	42.0	13	18 W10888	MAB anti-HBsAg bin
35	39.5	42.0	255	20 W88397	Mouse neuro-growth
36	39.5	42.0	275	21 Y52141	Mouse TANGO 125 (T
37	39.5	42.0	278	20 W88392	Mouse neuro-growth
38	39.5	42.0	302	15 R55793	Herpesvirus salmir
39	39.5	42.0	302	18 W26320	Herpesvirus secret
40	39.5	42.0	360	15 R55792	Herpesvirus salmir
41	39.5	42.0	360	18 W26319	Herpesvirus membra
42	39	41.5	15	18 W52659	Peptide which bind
43	39	41.5	15	18 W52663	Peptide which bind
44	39	41.5	15	18 W52643	Peptide which bind
45	39	41.5	15	18 W52648	Peptide which bind

ALIGNMENTS

RESULT 1

W12695	W12695 standard; Protein; 541 AA.
ID	W12695;
AC	W12695;
XX	
XX	
DT	31-MAY-1997 (first entry)
XX	
DE	G-protein parathyroid hormone receptor HLTGDG74.
XX	
XX	G-protein parathyroid hormone receptor; HLTGDG74; parathormone; PTH;
KW	calcium; signal transduction; agonist; antagonist; hypocalcaemia;
KW	hyperphosphataemia; hypoparathyroidism; chronic tetany;
KW	osteoporosis; hypercalcaemia; hypoparathyroidism; hypophosphataemia;
KW	kidney stone; nephrolithiasis; therapy; diagnosis.
XX	
XX	Homo sapiens.
XX	
PN	W09639433-A1.
XX	
PD	12-DEC-1996.
XX	
XX	05-JUN-1995; 95WO-US07085.
XX	
PR	05-JUN-1995; 95WO-US07085.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Li Y, Rosen CA, Ruben SM, Soppet DR;
XX	
XX	WPI; 1997-043068/04.
DR	
DR	N-PSDB; T59619.
XX	
PT	Human G-protein parathyroid hormone receptor, HLTGDG74 - used to
PT	identify (ant)agonists, used in the treatment of hypo- or

THIS PAGE BLANK (USPTO)

RP SEQUENCE FROM N.A.
RA Mallet L., Jacquet M.;
RT "Cloning and sequencing of a Candida albicans DNA fragment containing
RT the adenylate cyclase gene which complements CYR1 mutant as well as
RT six other genes including homologs of CHS6 and SAP185.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ250310; CAB59913.1; -.
KW Hypothetical protein.
SQ SEQUENCE 761 AA; 87454 MW; 45B3AF011E57DD0E CRC64;

Query Match 55.6%; Score 45; DB 3; Length 761;
Best Local Similarity 58.3%; Pred. No. 17;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NGEVQAEVKKMW 14
|||:|:|:
Db 400 NGLIASIRKKW 411

RESULT 13
Q9UPI1 PRELIMINARY; PRT; 82 AA.
AC Q9UPI1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE GIPR_HUMAN (FRAGMENT).
GN GIPR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K.J., Gordon L., Dias J., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala N., Terry A., Garnes J.,
RA Danganan L., Erlar A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Atlix C., Andreise T., Trankheim M., Amico-Keller G., Coefield J.,
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Carrano A.V.;
RT "Sequence analysis of a 1.9 Mb region in 19ql3.2 between APOE and
RT D19S412";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007191; AAD22674.1; -.
DR INTERPRO; IPR001749; -.
DR PRINTS; PR01129; GIPRECEPTOR.
FT NON_TER 1
SQ SEQUENCE 82 AA; 9101 MW; DF6E1CDE8459300C CRC64;

Query Match 54.3%; Score 44; DB 4; Length 82;
Best Local Similarity 50.0%; Pred. No. 2.7;
Matches 8; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

Qy 1 YC--NGEVQAEVKKMW 14
|||:|:|:
Db 8 YCFINKEVQSEIRRGW 23

RESULT 14
O49371 PRELIMINARY; PRT; 764 AA.
AC O49371;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-1998 (TrEMBLrel. 06, Last annotation update)
DE HYPOTHETICAL 88.1 KDA PROTEIN.
GN F1046.170.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Rosidae; eurosids II; Brassicales;

OC Brassicaceae; Arabidopsis.
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Weichselgartner M., Fartmann B., Granderath K., Dauner D.,
RA Herzl A., Neumann S., Hoheisel J., Mewes H.W., Mayer K., Schueller C.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL021811; CAA16971.1; -.
KW Hypothetical protein.
SQ SEQUENCE 764 AA; 88067 MW; 2F95FEB1850439F9 CRC64;

Query Match 54.3%; Score 44; DB 10; Length 764;
Best Local Similarity 53.8%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CNGEVQAEVKKMW 14
|||:|:|:
Db 727 CLGEVKEVVRLLW 739

RESULT 15
Q9PUK1 PRELIMINARY; PRT; 492 AA.
ID Q9PUK1
AC Q9PUK1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE GLUCAGON RECEPTOR.
GN GLUR.
OS Hoplobatrachus rugulosus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Rana.
RN [1]
RP SEQUENCE FROM N.A.
RA Ngan E.S.W., Chow L.S.N., Tse D.L.Y., Du X., Wei Y., Mojssov S.,
RA Chow B.K.C.;
RT "Functional studies of a glucagon receptor isolated from frog Rana
RT tigrina rugulosa: implications on the molecular evolution of glucagon
RT receptors in vertebrates";
RL FERS Lett. 0:0-0(1999).
DR EMBL; AF179589; AAD55777.1; -.
DR INTERPRO; IPR000832; -.
DR INTERPRO; IPR001749; -.
DR PFAM; PF00002; 7tm_2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PRINTS; PR01129; GIPRECEPTOR.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
KW Receptor.
SQ SEQUENCE 492 AA; 57150 MW; 32402A201CE87F5E CRC64;

Query Match 53.1%; Score 43; DB 13; Length 492;
Best Local Similarity 56.2%; Pred. No. 23;
Matches 9; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

Qy 1 YC--NGEVQAEVKKMW 14
|||:|:|:
Db 392 YCFVNKEVQSELLKKW 407

Search completed: November 8, 2000, 08:56:14
Job time: 525 sec

```

DR PRINTS; PRO1154; VIP1RECEPTOR.
DR PRINTS; PRO1156; PACAPRECEPTOR.
DR PROSITE; PS00649; G_PROTEIN_RECEPT_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEPT_F2_2; 1.
KW Receptor.
SQ SEQUENCE 459 AA; 52094 MW; C0C3A9AE1ADF611D CRC64;

Query Match 67.9%; Score 55; DB 11; Length 459;
Best Local Similarity 62.5%; Pred. No. 0.21;
Matches 10; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

QY 1 YC--NGEVOAEVKKMW 14
   || |||||:|:|:|
Db 390 YCFLNGEVOAEELRRKW 405

RESULT 9
ID Q920W0 PRELIMINARY; PRT; 550 AA.
AC Q920W0;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE GLUCAGON-LIKE PEPTIDE-2 RECEPTOR PRECURSOR.
GN GLP2R.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=HYPOTHALAMUS;
RX MEDLINE; 99145591.
RA Munroe D.G., Gupta A.K., Kooshesh F., Vyas T.B., Rizkalla G., Wang H.,
RA Demchyshyn L., Yang Z.-H., Kamboj R.K., Chen H., McCallum K.,
RA Sumner-Smith M., Drucker D.J., Crivici A.;
RT "Prototypic G protein-coupled receptor for the intestinotrophic factor
RT glucagon-like peptide 2.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:1569-1574(1999).
DR EMBL; AF105367; AAD16895.1; -
DR INTERPRO; IPR000832; -
DR INTERPRO; IPR001879; -
DR PFAM; PF00002; 7tm_2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PROSITE; PS00649; G_PROTEIN_RECEPT_F2_1; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 550 AA; 63102 MW; 22E269F81E25226 CRC64;

Query Match 64.2%; Score 52; DB 11; Length 550;
Best Local Similarity 57.1%; Pred. No. 0.8;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 YCNGEVOAEVKKMW 14
   : |||||:|:|:|
Db 438 FANGEVRAELRRKW 451

RESULT 10
ID Q95838 PRELIMINARY; PRT; 553 AA.
AC Q95838;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE GLUCAGON-LIKE PEPTIDE-2 RECEPTOR PRECURSOR.
GN GLP2R.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=STOMACH;

Query Match 64.2%; Score 52; DB 11; Length 550;
Best Local Similarity 57.1%; Pred. No. 0.8;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 YCNGEVOAEVKKMW 14
   : |||||:|:|:|
Db 438 FANGEVRAELRRKW 451

RESULT 11
ID Q73768 PRELIMINARY; PRT; 438 AA.
AC Q73768;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE GROWTH-HORMONE-RELEASING HORMONE-LIKE PEPTIDE RECEPTOR.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprinidae; Cyprinidae; Carassius.
RN [1]
RP SEQUENCE FROM N.A.
RA Chan K.W., Yu K.L., Rivier J., Chow B.K.C.;
RL Neuroendocrinology 0:0-0(1998).
DR EMBL; AF048819; AAC15698.1; -
DR INTERPRO; IPR000832; -
DR INTERPRO; IPR001879; -
DR PFAM; PF00002; 7tm_2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PROSITE; PS00650; G_PROTEIN_RECEPT_F2_2; UNKNOWN_1.
SQ SEQUENCE 438 AA; 50305 MW; C3516A52D692C339 CRC64;

Query Match 58.6%; Score 47.5; DB 13; Length 438;
Best Local Similarity 58.8%; Pred. No. 3.6;
Matches 10; Conservative 4; Mismatches 0; Indels 3; Gaps 2;

QY 1 YC--NGEVOAEVK-KMW 14
   || |||||:|:|
Db 383 YCFLNGDVQAEELRRKW 399

RESULT 12
ID Q9UVJ1 PRELIMINARY; PRT; 761 AA.
AC Q9UVJ1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE HYPOTHETICAL 87.5 KDA PROTEIN.
GN CHS6.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC anamorphic Saccharomycetales; Candida.
RN [1]

```

```
DR INTERPRO: IPR000832; -
DR INTERPRO: IPR001571; -
DR INTERPRO: IPR001771; -
DR INTERPRO: IPR001879; -
DR INTERPRO: IPR002285; -
DR PFAM: PF00002; 7tm_2; 1.
DR PRINTS: PR00249; GPCRSECRETIN.
DR PRINTS: PR00491; VASOACTIVEIPR.
DR PRINTS: PR01154; VIP1RECEPTOR.
DR PRINTS: PR01156; PACAPRECEPTOR.
DR PROSITE: PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE: PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 444 AA; 50955 MW; 883B25B729314C4C CRC64;

Query Match 71.68; Score 58; DB 13; Length 444;
Best Local Similarity 68.8; Pred. No. 0.063;
Matches 11; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 1 YC--NGEVOAEVKKMW 14
II |IIIIII:|:|
DB 375 YCFLNGEVOAEVKKRW 390

RESULT 6
O73769 PRELIMINARY; PRT; 465 AA.
AC O73769;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE PITUITARY ADENYLATE CYCLASE ACTIVATING POLYPEPTIDE TYPE 1 RECEPTOR
DE PRECURSOR.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprinidae; Cyprinidae; Carassius.
[1]
RP SEQUENCE FROM N.A.
RA Wong A.O.L., Leung M.Y., Shea W.L.C., Chang J.P., Chow B.K.C.;
RL Submitted (Feb-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF048820; AAC15699.1; -
DR INTERPRO: IPR000832; -
DR INTERPRO: IPR001879; -
DR PFAM: PF00002; 7tm_2; 1.
DR PRINTS: PR00249; GPCRSECRETIN.
DR PROSITE: PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE: PS00650; G_PROTEIN_RECEP_F2_2; 1.
SQ SEQUENCE 465 AA; 53459 MW; 9779A95EDBFD1DC1 CRC64;

Query Match 70.4%; Score 57; DB 13; Length 465;
Best Local Similarity 62.5%; Pred. No. 0.098;
Matches 10; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

QY 1 YC--NGEVOAEVKKMW 14
II |IIIIII:|:|
DB 392 YCFLNGEVOAEVKKRW 407

RESULT 7
O9PTK1 PRELIMINARY; PRT; 465 AA.
AC O9PTK1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE PITUITARY ADENYLATE CYCLASE-ACTIVATING POLYPEPTIDE TYPE I RECEPTOR.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Hu Z., Lelievre V., Chao A., Zhou X., Waschek J.A.;
RT "Characterization and mRNA distribution of a cloned pituitary
RT adenylate cyclase-activating polypeptide type I receptor in the
RT Xenopus brain.";
RL Endocrinology 0:0-0(2000).
DR EMBL: AF187878; AAF16939.1; -
DR INTERPRO: IPR000832; -
DR INTERPRO: IPR001771; -
DR INTERPRO: IPR001879; -
DR INTERPRO: IPR002285; -
DR PFAM: PF00002; 7tm_2; 1.
DR PRINTS: PR00249; GPCRSECRETIN.
DR PRINTS: PR01154; VIP1RECEPTOR.
DR PRINTS: PR01156; PACAPRECEPTOR.
DR PROSITE: PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE: PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 465 AA; 53424 MW; D16C413523EFAE43 CRC64;

Query Match 70.4%; Score 57; DB 13; Length 465;
Best Local Similarity 62.5%; Pred. No. 0.098;
Matches 10; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

QY 1 YC--NGEVOAEVKKMW 14
II |IIIIII:|:|
DB 397 YCFLNGEVOAEVKKRW 412

RESULT 8
O9RT8 PRELIMINARY; PRT; 459 AA.
AC O9RT8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE VASOACTIVE INTESTINAL POLYPEPTIDE 1 (VPAC1) RECEPTOR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=129SVJ;
RA Hashimoto H., Nishino A., Shintani N., Hagiwara N., Copeland N.G.,
RA Jenkins N.A., Yamamoto K., Matsuda T., Ishihara T., Nagata S.,
RA Baba A.;
RT "Genomic organization and chromosomal location of the mouse vasoactive
RT intestinal polypeptide 1 (VPAC1) receptor.";
RL Genomics 58:90-93(1999).
DR EMBL: AB022860; BAA81896.1; -
DR EMBL: AB022848; BAA81896.1; JOINED.
DR EMBL: AB022849; BAA81896.1; JOINED.
DR EMBL: AB022850; BAA81896.1; JOINED.
DR EMBL: AB022851; BAA81896.1; JOINED.
DR EMBL: AB022852; BAA81896.1; JOINED.
DR EMBL: AB022853; BAA81896.1; JOINED.
DR EMBL: AB022854; BAA81896.1; JOINED.
DR EMBL: AB022855; BAA81896.1; JOINED.
DR EMBL: AB022856; BAA81896.1; JOINED.
DR EMBL: AB022857; BAA81896.1; JOINED.
DR EMBL: AB022858; BAA81896.1; JOINED.
DR EMBL: AB022859; BAA81896.1; JOINED.
DR INTERPRO: IPR000832; -
DR INTERPRO: IPR001571; -
DR INTERPRO: IPR001771; -
DR INTERPRO: IPR001879; -
DR INTERPRO: IPR002285; -
DR PFAM: PF00002; 7tm_2; 1.
DR PRINTS: PR00249; GPCRSECRETIN.
DR PRINTS; PR00491; VASOACTIVEIPR.
```

```

RESULT 2
Q9PVD3 ID Q9PVD3 PRELIMINARY; PRT; 536 AA.
AC Q9PVD3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE PARATHYROID HORMONE RECEPTOR PTHLR.
OS Brachydanio rerio (zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprinidae; Rasbora; Danio.
RN [1]
RP SEQUENCE FROM N.A.
RA Rubin D.A., Juppner H.;
RT "Zebrafish Express the Common Parathyroid Hormone/Parathyroid Hormone-
related Peptide Receptor (PTHLR) and a Novel Receptor (PTH3R) That Is
preferentially Activated by Mammalian and Fuguish Parathyroid
Hormone-related Peptide."
RL J. Biol. Chem. 274:28185-28190(1999).
DR EMBL; AF132084; AAF01265.1; -.
DR INTERPRO; IPR000832; -.
DR INTERPRO; IPR001879; -.
DR INTERPRO; IPR002170; -.
DR PFAM; PF00002; 7tm_2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PROSITE; PS00393; PTRHORMONER.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 536 AA; 61454 MW; 653B0A56DB2C8FA9 CRC64;

Query Match 86.4%; Score 70; DB 13; Length 536;
Best Local Similarity 78.6%; Pred. No. 0.00073;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YCNGEVAQAEVKKMW 14
:|||||:|
Db 415 FCNGEVAQAEIKKAW 428

RESULT 3
Q9PWB7 ID Q9PWB7 PRELIMINARY; PRT; 575 AA.
AC Q9PWB7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE PARATHYROID HORMONE TYPE-2 RECEPTOR PRECURSOR.
OS Brachydanio rerio (zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprinidae; Rasbora; Danio.
RN [1]
RP SEQUENCE FROM N.A.
RA Rubin D.A., Hellman P., Zon L.I., Lobb C.J., Bergwitz C., Juppner H.;
RT "A G protein-coupled receptor from zebrafish is activated by human
parathyroid hormone and not by human or teleost parathyroid hormone-
related peptide. Implications for the evolutionary conservation of
calcium-regulating peptide hormones."
RL J. Biol. Chem. 274:23035-23042(1999).
DR EMBL; AF132082; AAD51908.1; -.
DR INTERPRO; IPR000832; -.
DR INTERPRO; IPR001879; -.
DR PFAM; PF00002; 7tm_2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 575 AA; 64244 MW; 888F1C4DDB3A14DC CRC64;

```

```

Query Match 86.4%; Score 70; DB 13; Length 575;
Best Local Similarity 78.6%; Pred. No. 0.00079;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YCNGEVAQAEVKKMW 14
:|||||:|
Db 442 YCNGEVAQAEIKKWTW 455

RESULT 4
Q9TU31 ID Q9TU31 PRELIMINARY; PRT; 595 AA.
AC Q9TU31;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE PARATHYROID HORMONE RECEPTOR-1.
GN PTHL.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN [1]
RP SEQUENCE FROM N.A.
RA Smock S.L., Vogt G.A., Castleberry T.A., Lu B., Owen T.A.;
RT "Molecular cloning and functional characterization of the canine
parathyroid hormone receptor-1 (PTH1)."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF167095; AAD55938.1; -.
DR INTERPRO; IPR000832; -.
DR INTERPRO; IPR001879; -.
DR INTERPRO; IPR002170; -.
DR PFAM; PF00002; 7tm_2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PROSITE; PS00393; PTRHORMONER.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 595 AA; 66308 MW; 09568ECF38D4D258 CRC64;

Query Match 86.4%; Score 70; DB 6; Length 595;
Best Local Similarity 78.6%; Pred. No. 0.00081;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YCNGEVAQAEVKKMW 14
:|||||:|
Db 460 FCNGEVAQAEIKKSW 473

RESULT 5
Q9YHC6 ID Q9YHC6 PRELIMINARY; PRT; 444 AA.
AC Q9YHC6;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE VASOACTIVE INTESTINAL PEPTIDE/PITUITARY ADENYLATE CYCLASE ACTIVATING
POLYPEPTIDE RECEPTOR.
OS Rana ridibunda (Laughing frog) (Marsh frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.
RN [1]
RP SEQUENCE FROM N.A.
RA Alexandre D., Anouar Y.;
RT "A cloned frog VIP/PACAP receptor exhibits pharmacological and tissue
distribution characteristics of both VPAC1 and VPAC2 receptors in
mammals."
RL Endocrinology 0:0-0(1999).
DR EMBL; AF100644; AAD03602.1; -.

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: November 8, 2000, 08:56:12 ; Search time 152.43 Seconds
(without alignments)
8.576 Million cell updates/sec

Title: US-09-236-468A-2_COPY_415_428
Perfect score: 81
Sequence: 1 YCNGEQVAEVKKMW 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 297973

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB ID	Description
1	71	87.7	542	13 Q9PVD2	Q9pvd2 brachydanio
2	70	86.4	536	13 Q9PVD3	Q9pvd3 brachydanio
3	70	86.4	575	13 Q9PWB7	Q9pwb7 brachydanio
4	70	86.4	595	6 Q9TU31	Q9tcu31 canis famli
5	58	71.6	444	13 Q9YHC6	Q9yhc6 rana ridibu
6	57	70.4	465	13 Q9P769	Q9p769 carassius a
7	57	70.4	465	13 Q9PTK1	Q9ptk1 xenopus lae
8	55	67.9	459	11 Q9R1T8	Q9rlt8 mus musculu
9	52	64.2	550	11 Q9Z0W0	Q9z0w0 rattus norv
10	52	64.2	553	4 Q9S838	Q9s838 homo sapien
11	47.5	58.6	438	13 Q93768	Q93768 carassius a
12	45	55.6	761	3 Q9UVJ1	Q9uvj1 candida alb
13	44	54.3	82	4 Q9UP11	Q9up11 homo sapien
14	44	54.3	764	10 Q49371	Q49371 arabidopsis
15	43	53.1	492	13 Q9PUK1	Q9puk1 hoplobatrac
16	43	53.1	568	3 Q74639	Q74639 gibberella
17	42.5	52.5	195	2 Q51113	Q51113 neisseria m
18	42	51.9	410	2 P72841	P72841 synechocyst
19	42	51.9	463	11 Q9WUP2	Q9wup2 mus musculu

20	42	51.9	463	11 Q9R1W5	Q9rlw5 mus musculu
21	42	51.9	463	11 Q9QXH8	Q9qxh8 mus musculu
22	42	51.9	682	5 Q9VAS2	Q9vas2 drosophila
23	41	50.6	102	2 Q9RLO9	Q9rlq9 mycobacteri
24	41	50.6	251	12 Q67884	Q67884 groundnut r
25	41	50.6	252	10 Q9X125	Q9x125 arabidopsis
26	41	50.6	413	2 P96288	P96288 mycobacteri
27	40	49.4	167	2 P71299	P71299 escherichia
28	40	49.4	300	1 Q28459	Q28459 archaeoglob
29	40	49.4	434	2 Q52748	Q52748 ruminococcu
30	40	49.4	494	5 Q44515	Q44515 caenorhabdi
31	40	49.4	1146	5 Q17482	Q17482 caenorhabdi
32	39	48.1	145	5 Q9Y1Z2	Q9y1z2 hydra magni
33	39	48.1	235	2 Q9XC14	Q9xc14 shigella fl
34	39	48.1	439	11 Q9WU99	Q9wu99 rattus norv
35	39	48.1	441	6 Q9TUJ0	Q9tuj0 bos taurus
36	39	48.1	501	4 Q12874	Q12874 homo sapien
37	39	48.1	501	4 Q15460	Q15460 homo sapien
38	39	48.1	1002	5 Q45247	Q45247 caenorhabdi
39	38	46.9	219	2 Q56931	Q56931 versinia en
40	38	46.9	343	1 Q58932	Q58932 methanococc
41	38	46.9	343	5 Q9XXA5	Q9xxa5 caenorhabdi
42	38	46.9	347	4 Q43866	Q43866 homo sapien
43	38	46.9	424	10 Q9SKU5	Q9sku5 arabidopsis
44	38	46.9	444	5 Q26861	Q26861 trypanosoma
45	38	46.9	466	10 Q9SS64	Q9ss64 arabidopsis

ALIGNMENTS

RESULT	1
Q9PVD2	
ID	Q9PVD2 PRELIMINARY; PRT; 542 AA.
AC	Q9PVD2;
DT	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE	PARATHYROID HORMONE-RELATED PROTEIN RECEPTOR PTH3R.
OS	Brachydanio rerio (zebrafish) (zebra danio).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC	Cypriniformes; Cyprinidae; Cyprininae; Rasbora; Danio.
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Rubin D.A., Jueppner H.;
RT	"Zebrafish Express the Common Parathyroid Hormone/Parathyroid Hormone-
RT	related Peptide Receptor (PTH1R) and a Novel Receptor (PTH3R) That Is
RT	Preferentially Activated by Mammalian and Fugufish Parathyroid
RT	Hormone-related Peptide."
RL	J. Biol. Chem. 274:28185-28190(1999).
DR	EMBL; AF132085; AAF01266.2; -.
DR	INTERPRO; IPR000832; -.
DR	INTERPRO; IPR001879; -.
DR	INTERPRO; IPR002170; -.
DR	PFAM; PF00002; 7tm_2; 1.
DR	PRINTS; PR00249; GPCRSECRETIN.
DR	PRINTS; PR00393; PTHRMONER.
DR	PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR	PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW	Receptor.
SQ	SEQUENCE 542 AA; 61438 MW; 08688658E2727303 CRC64;

Query Match 87.7%; Score 71; DB 13; Length 542;
Best Local Similarity 85.7%; Pred. No. 0.0005;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCNGEQVAEVKKMW 14
:|||||
DB 400 FCNGEQVAEVKKAW 413

THIS PAGE BLANK (USPTO)

FT CONFLICT 124 124 G -> A (IN REF. 1).
FT CONFLICT 210 210 A -> P (IN REF. 2).
FT CONFLICT 308 308 I -> F (IN REF. 3).
FT CONFLICT 333 333 E -> Q (IN REF. 3).
FT CONFLICT 377 377 G -> A (IN REF. 1).
SQ SEQUENCE 440 AA; 50206 MW; E22CDD0EE7C0ACCL CRC64;

Query Match 67.9%; Score 55; DB 1; Length 440;
Best Local Similarity 68.8%; Pred. No. 0.05;
Matches 11; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Qy 1 YC--NGEVOAEVKKMW 14
|| ||||| ||:| |
Db 388 YCFLNGEVLQEVQKKW 403

RESULT 15
SCRC_RABIT
ID SCRC_RABIT STANDARD; PRT; 445 AA.
AC O46502;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SECRETIN RECEPTOR PRECURSOR (SCT-R).
GN SCTR.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98366112.
RA Svoboda M., Tastenoy M., de Neef P., Delporte C., Waelbroeck M.,
RA Robberecht P.;
RT "Molecular cloning and in vitro properties of the recombinant rabbit
secretin receptor.";
RL Peptides 19:1055-1062(1998).
CC -!- FUNCTION: THIS IS A RECEPTOR FOR SECRETIN. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYL
CC CYCLASE.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF025411; AAC32767.1; -.
DR INTERPRO; IPR000832; -.
DR INTERPRO; IPR001771; -.
DR INTERPRO; IPR002144; -.
DR PFAM; PF00002; 7tm.2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PRINTS; PR00490; SECRETINR.
DR PRINTS; PR01154; VIPRECEPTOR.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 21 POTENTIAL
FT CHAIN 22 445 SECRETIN RECEPTOR.
FT DOMAIN 22 139 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 140 163 1 (POTENTIAL).
FT DOMAIN 164 170 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 171 190 2 (POTENTIAL).
FT DOMAIN 191 212 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 213 236 3 (POTENTIAL).
FT DOMAIN 237 250 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 251 272 4 (POTENTIAL).
FT DOMAIN 273 290 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 291 313 5 (POTENTIAL).
FT DOMAIN 314 339 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 340 358 6 (POTENTIAL).
FT DOMAIN 359 365 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 366 388 7 (POTENTIAL).
FT DOMAIN 389 445 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 445 AA; 50495 MW; 31C4169CB099F194 CRC64;

Query Match 67.9%; Score 55; DB 1; Length 445;
Best Local Similarity 68.8%; Pred. No. 0.051;
Matches 11; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Qy 1 YC--NGEVOAEVKKMW 14
|| ||||| ||:| |
Db 384 YCFLNGEVLQEVQKKW 399

Search completed: November 8, 2000, 09:03:52
Job time: 862 sec

DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE VASOACTIVE INTESTINAL POLYPEPTIDE RECEPTOR (VIP-R) (VIP RECEPTOR)
 OS Meleagris gallopavo (Common Turkey).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Meleagridae; Meleagris.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SMALL INTESTINE;
 RX MEDLINE; 96206340.
 RA Xu M., Proudman J.A., Pitts G.R., Wong E.A., Foster D.N.,
 el Halawani M.E.;
 RT "Vasoactive intestinal peptide stimulates prolactin mRNA expression
 in turkey pituitary cells: effects of dopaminergic drugs.";
 RL Proc. Soc. Exp. Biol. Med. 212:52-62(1996).
 CC -!- FUNCTION: THIS IS A RECEPTOR FOR VIP. THE ACTIVITY OF THIS
 CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL
 CC CYCLASE.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U31991; AAA9740.1; -;
 DR GCRDB; GCR_1749; -;
 DR INTERPRO; IPR000832; -;
 DR PFAM; PF00002; 7tm_2; 1.
 DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; PARTIAL.
 DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT NON_TER 1
 FT DOMAIN <1 18 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 19 42 3 (POTENTIAL).
 FT DOMAIN 43 56 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 57 78 4 (POTENTIAL).
 FT DOMAIN 79 95 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 96 119 5 (POTENTIAL).
 FT DOMAIN 120 144 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 145 164 6 (POTENTIAL).
 FT DOMAIN 165 176 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 177 196 7 (POTENTIAL).
 FT DOMAIN 197 260 CYTOPLASMIC (POTENTIAL).
 SQ SEQUENCE 260 AA; 30358 MW; 15761E6AB5B23D5A CRC64;

 Query Match 71.6%; Score 58; DB 1; Length 260;
 Best Local Similarity 68.8%; Pred. No. 0.0091;
 Matches 11; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

 Qy 1 YC--NGEVQAEVKKW 14
 Db 191 YCFLNGEVQAEVKKW 206
 || |||||:|:|

 RESULT 14
 SCRC_HUMAN STANDARD; PRT; 440 AA.
 AC P47872; Q13213; Q12961;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE SECRETIN RECEPTOR PRECURSOR (SCT-R).
 GN SCTR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=PANCREAS;
 RX MEDLINE; 95169147.
 RA Jlang S., Ulrich C.D.;
 RT "Molecular cloning and functional expression of a human pancreatic
 secretin receptor.";
 RL Biochem. Biophys. Res. Commun. 207:883-890(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PANCREAS;
 RX MEDLINE; 95336443.
 RA Chow B.K.-C.;
 RT "Molecular cloning and functional characterization of a human
 secretin receptor.";
 RL Biochem. Biophys. Res. Commun. 212:204-211(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RX MEDLINE; 95214632.
 RA Patel D.R., Kong Y., Sreedharan S.P.;
 RT "Molecular cloning and expression of a human secretin receptor.";
 RL Mol. Pharmacol. 47:467-473(1995).
 CC -!- FUNCTION: THIS IS A RECEPTOR FOR SECRETIN. THE ACTIVITY OF THIS
 CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL
 CC CYCLASE.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U20178; AAC50106.1; -;
 DR EMBL; U28281; AAA87556.1; -;
 DR EMBL; U13989; AAA64949.1; -;
 DR GCRDB; GCR_1995; -;
 DR GCRDB; GCR_2016; -;
 DR GCRDB; GCR_2033; -;
 DR MIN; 182098; -;
 DR INTERPRO; IPR000832; -;
 DR INTERPRO; IPR002144; -;
 DR PFAM; PF00002; 7tm_2; 1.
 DR PRINTS; PR00249; GPCRSECRETIN.
 DR PRINTS; PR00490; SECRETINR.
 DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
 DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 440 SECRETIN RECEPTOR.
 FT DOMAIN 23 143 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 144 167 1 (POTENTIAL).
 FT DOMAIN 168 174 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 175 194 2 (POTENTIAL).
 FT DOMAIN 195 216 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 217 240 3 (POTENTIAL).
 FT DOMAIN 241 254 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 255 276 4 (POTENTIAL).
 FT DOMAIN 277 294 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 295 317 5 (POTENTIAL).
 FT DOMAIN 318 343 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 344 362 6 (POTENTIAL).
 FT DOMAIN 363 369 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 370 392 7 (POTENTIAL).
 FT DOMAIN 393 440 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).

adenylate cyclase-activating polypeptide type I receptor." ;
Proc. Natl. Acad. Sci. U.S.A. 90:6345-6349(1993).

[2]
SEQUENCE FROM N.A.

TISSUE-BRAIN;
MEDLINE; 93326107.
Hosoya M., Onda H., Ogi K., Masuda Y., Miyamoto Y., Ohtaki T.,
Okazaki H., Arimura A., Fujino M.;

"Molecular cloning and functional expression of the receptor for pituitary adenylate cyclase activating peptide (PACAP).";
Biochem. Biophys. Res. Commun. 194:133-143(1993).

[3] SEQUENCE FROM N.A.
TISSUE-PITUITARY:
MEDLINE: 93357025.
Hashimoto H., Ishihara T., Shigemoto R., Mori K., Nagata S.:
"Molecular cloning and tissue distribution of a receptor for
pituitary adenylate cyclase-activating polypeptide.";
Neuron 11:333-342(1993).

[4] SEQUENCE FROM N.A.
STRAIN-WISTAR.
MEDLINE: 93382505.
Spengler D., Waeber C., Pantaloni C., Holsboer F., Bockaert J.,
Seeburg P.H., Journot L.;
"Differential signal transduction by five splice variants of the
PACAP receptor.";
Nature 365:170-175(1993).

[5] SEQUENCE FROM N.A.
STRAIN-SPRAGUE-DAWLEY; TISSUE-OLFACTORY BULB;
MEDLINE; 93359075;
Morrow J.A., Lutz E.M., West K.M., Fink G., Harmar A.J.;
"Molecular cloning and expression of a cDNA encoding a receptor for
pituitary adenylate cyclase activating polypeptide (PACAP).";
FEBS Lett. 329:95-105(1993).

[6] SEQUENCE OF 115-523 FROM N.A. STRAIN-WISTAR;
Svoboda M., Ciccarelli E., Tastency M., Christophe J.;
Submitted (May-1993) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: THIS IS A RECEPTOR FOR PACAP-27 AND PACAP-38. THE
ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
ADENYLYL CYCLASE. MAY REGULATE THE RELEASE OF ADRENOCORTICOTROPIN,
LUTEINIZING HORMONE, GROWTH HORMONE, PROLACTIN, EPINEPHRINE, AND
CATECHOLAMINE. MAY PLAY A ROLE IN SPERMATOGENESIS AND SPERM
MOTILITY. CAUSES SMOOTH MUSCLE RELAXATION AND SECRETION IN THE
GASTROINTESTINAL TRACT.
-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
-!- ALTERNATIVE PRODUCTS: FIVE ISOFORMS; HIP-HOP1 (SHOWN HERE), HOP1,
HOP2, HIP AND PACAP-R; ARE PRODUCED BY ALTERNATIVE SPLICING.
-!- TISSUE SPECIFICITY: HYPOTHALAMUS, ANTERIOR PITUITARY, ADRENAL
MEDULLA, TESTICULAR GERM CELLS.
-!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed, usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; D16465; BAA03932.1; -;
EMBL; L16680; AAA41792.1; -;
EMBL; D14908; BAA03608.1; -;
EMBL; D14909; BAA03609.1; -;
EMBL; D23272; CA80810.1; -;
EMBL; D23273; CA80811.1; -;
EMBL; D23274; CA80812.1; -;
EMBL; D23275; CA80813.1; -;
EMBL; D23279; CA80817.1; -;

[illegible]

Query Match 74.1%; Score 60; DB 1; Length 523;
Best Local Similarity 68.8%; Pred. No. 0.0084;
Matches 11: Conservative 2: Mismatches 1: Indels

Qy 1 YC--NGEVOAEVKKMW 14
 || |||||:|:
 Db 455 YCFNLGEVOAEIKRKW 470

RESULT 13					
VIPR_MELGA					
ID	VIPR_MELGA	STANDARD;	PRT;	260 AA.	
AC	Q31085;				
DT	01-NOV-1997	(Rel. 35, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			

```

DR PROSITE; PS00650; G_PROTEIN_RECP_F2_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 496 PITUITARY ADENYLATE CYCLASE ACTIVATING
FT POLYPEPTIDE TYPE I RECEPTOR.
FT EXTRACELLULAR (POTENTIAL).
FT DOMAIN 21 155 1 (POTENTIAL).
FT TRANSMEM 156 178 2 (POTENTIAL).
FT DOMAIN 179 186 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 187 205 2 (POTENTIAL).
FT DOMAIN 206 227 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 228 253 3 (POTENTIAL).
FT DOMAIN 254 268 4 (POTENTIAL).
FT TRANSMEM 269 291 5 (POTENTIAL).
FT DOMAIN 292 309 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 310 332 6 (POTENTIAL).
FT DOMAIN 333 378 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 379 399 7 (POTENTIAL).
FT DOMAIN 400 413 8 (POTENTIAL).
FT TRANSMEM 414 433 9 (POTENTIAL).
FT DOMAIN 434 496 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 117 117 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 496 AA; 56639 MW; 47B5D51D4209060A CRC64;

Query Match 74.1%; Score 60; DB 1; Length 496;
Best Local Similarity 68.8%; Pred. No. 0.0079;
Matches 11; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 1 YC--NGEVQAEVKKMW 14
II IIIIIII:I
DB 428 YCFLNGEVQAEIKRKW 443

RESULT 11
ID PACR_BOVIN STANDARD; PRT; 513 AA.
AC Q29627;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PITUITARY ADENYLATE CYCLASE ACTIVATING POLYPEPTIDE TYPE I RECEPTOR
DE PRECURSOR (PACAP TYPE I RECEPTOR) (PACAP-R-1).
GN ADCYAP1R1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=BRAIN;
RA Miyamoto Y., Habata Y., Ohtaki T., Masuda Y., Ogi K., Onda H.,
RA Fujino M.;
RT "Cloning and expression of a complementary DNA encoding the bovine
RT receptor for pituitary adenylate cyclase-activating polypeptide
RT (PACAP).";
RL Blochim. Biophys. Acta 1218:297-307(1994).
CC -1- FUNCTION: THIS IS A RECEPTOR FOR PACAP-27 AND PACAP-38. THE
CC ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLATE CYCLASE. MAY REGULATE THE RELEASE OF ADRENOCORTICOTROPIN,
CC LUTEINIZING HORMONE, GROWTH HORMONE, PROLACTIN, EPINEPHRINE, AND
CC CATECHOLAMINE. MAY PLAY A ROLE IN SPERMATOGENESIS AND SPERM
CC MOTILITY. CAUSES SMOOTH MUSCLE RELAXATION AND SECRETION IN THE
CC GASTROINTESTINAL TRACT (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; A LONG FORM (SHOWN
CC HERE) AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D17290; BAA04122.1; -.
CC EMBL; D17290; BAA04123.1; -.
CC GCRDB; GCR_2658; -.
CC GCRDB; GCR_2659; -.
CC INTERPRO; IPR000832; -.
CC INTERPRO; IPR002285; -.
CC PFAM; PF00002; 7tm_2; 1.
CC PRINTS; PR00249; GPCRSECRETIN.
CC PRINTS; PR01156; PACAPRECEPTR.
CC PROSITE; PS00649; G_PROTEIN_RECP_F2_1; 1.
CC PROSITE; PS00650; G_PROTEIN_RECP_F2_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Alternative splicing.
FT SIGNAL 1 37 POTENTIAL.
FT CHAIN 38 513 PITUITARY ADENYLATE CYCLASE ACTIVATING
FT POLYPEPTIDE TYPE I RECEPTOR.
FT EXTRACELLULAR (POTENTIAL).
FT DOMAIN 38 172 1 (POTENTIAL).
FT TRANSMEM 173 195 2 (POTENTIAL).
FT DOMAIN 196 203 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 204 222 3 (POTENTIAL).
FT DOMAIN 223 244 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 245 270 4 (POTENTIAL).
FT DOMAIN 271 285 5 (POTENTIAL).
FT TRANSMEM 286 308 6 (POTENTIAL).
FT DOMAIN 309 326 7 (POTENTIAL).
FT TRANSMEM 327 349 8 (POTENTIAL).
FT DOMAIN 350 395 9 (POTENTIAL).
FT TRANSMEM 396 416 10 (POTENTIAL).
FT DOMAIN 417 430 11 (POTENTIAL).
FT TRANSMEM 431 450 12 (POTENTIAL).
FT DOMAIN 451 513 13 (POTENTIAL).
FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 366 393 MISSING (IN SHORT ISOFORM).
SQ SEQUENCE 513 AA; 58785 MW; 9A45753210CE9817 CRC64;

Query Match 74.1%; Score 60; DB 1; Length 513;
Best Local Similarity 68.8%; Pred. No. 0.0082;
Matches 11; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 1 YC--NGEVQAEVKKMW 14
II IIIIIII:I
DB 445 YCFLNGEVQAEIKRKW 460

RESULT 12
ID PACR_RAT STANDARD; PRT; 523 AA.
AC P32215; Q63414;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PITUITARY ADENYLATE CYCLASE ACTIVATING POLYPEPTIDE TYPE IA RECEPTOR
DE PRECURSOR (PACAP TYPE IA RECEPTOR) (PACAP-R-1).
GN ADCYAP1R1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=PANCREAS;
RX MEDLINE; 93317678.
RA Wank S.A., Pisseña J.R.;
RA "Molecular cloning and functional expression of the pituitary

```

```

ID AC PACR_HUMAN STANDARD; PRT; 468 AA.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PITUITARY ADENYLATE CYCLASE ACTIVATING POLYPEPTIDE TYPE I RECEPTOR
DE PRECURSOR (PACAP TYPE I RECEPTOR) (PACAP-R-1).
GN ADCYAP1R1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PITUITARY;
RX MEDLINE; 94071918.
RA Ogi K., Miyamoto Y., Masuda Y., Habata Y., Hosoya M., Ohtaki T.,
RA Masuo Y., Onda H., Fujino M.;
RT "Molecular cloning and functional expression of a cDNA encoding a
RT human pituitary adenylate cyclase activating polypeptide receptor.";
RL Biochem. Biophys. Res. Commun. 196:1511-1521(1993).
RN [2]
RP SEQUENCE OF 418-468 FROM N.A.
RC TISSUE-PLACENTA;
RX MEDLINE; 95154842.
RA Stoffel M., Espinosa R., Trabb J.B., le Beau M.M., Bell G.I.;
RT "Human type I pituitary adenylate cyclase activating polypeptide
RT receptor (ADCYAP1R); localization to chromosome band 7p14 and
RT integration into the cytogenetic, physical and genetic map of
RT chromosome 7";
RL Genomics 23:697-699(1994).
CC -!- FUNCTION: THIS IS A RECEPTOR FOR PACAP-27 AND PACAP-38. THE
CC ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLATE CYCLASE. MAY REGULATE THE RELEASE OF ADRENOCORTICOTROPIN,
CC LUTEINIZING HORMONE, GROWTH HORMONE, PROLACTIN, EPINEPHRINE, AND
CC CATECHOLAMINE. MAY PLAY A ROLE IN SPERMATOGENESIS AND SPERM
CC MOTILITY. CAUSES SMOOTH MUSCLE RELAXATION AND SECRETION IN THE
CC GASTROINTESTINAL TRACT.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: MOST ABUNDANT IN THE BRAIN, LOW EXPRESSION IN
CC THE LUNG, LIVER, THYMUS, SPLEEN, PANCREAS AND PLACENTA.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; A28104; CAA01920.1; -.
CC EMBL; AF061785; AAC31809.1; -.
CC GCRDB; GCR_0899; -.
CC MM; 102981; -.
CC INTERPRO; IPR000832; -.
CC INTERPRO; IPR002285; -.
CC PFAM; PF00002; 7tm_2; 1.
CC PRINTS; PR00249; GPCRSECRETIN.
CC PRINTS; PR01156; PACAPRECEPTR.
CC PROSITE; PS00649; G_PROTEIN_RECEPT_F2_1; 1.
CC PROSITE; PS00650; G_PROTEIN_RECEPT_F2_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
CC Spermatogenesis.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 468 PITUITARY ADENYLATE CYCLASE ACTIVATING
FT POLYPEPTIDE TYPE I RECEPTOR.
FT DOMAIN 21 155 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 156 178 1 (POTENTIAL).
FT DOMAIN 179 186 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 187 205 2 (POTENTIAL).
FT DOMAIN 206 227 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 228 253 3 (POTENTIAL).
FT DOMAIN 254 268 CYTOPLASMIC (POTENTIAL).

```

```

FT TRANSSEM 269 291 4 (POTENTIAL).
FT DOMAIN 292 309 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 310 332 5 (POTENTIAL).
FT DOMAIN 333 350 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 351 371 6 (POTENTIAL).
FT DOMAIN 372 385 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 386 405 7 (POTENTIAL).
FT DOMAIN 406 468 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 117 117 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 375 375 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 468 AA; 53313 MW; BB515B84E9F28977 CRC64;

Query Match 74.1%; Score 60; DB 1; Length 468;
Best Local Similarity 68.8%; Pred. No. 0.0075;
Matches 11; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 1 YC--NGEVOAEVKKMW 14
|| |||||:::|
Db 400 YCFLNGEVOAEIKRW 415

RESULT 10
PACR_MOUSE STANDARD; PRT; 496 AA.
AC P70205;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PITUITARY ADENYLATE CYCLASE ACTIVATING POLYPEPTIDE TYPE I RECEPTOR
DE PRECURSOR (PACAP TYPE I RECEPTOR) (PACAP-R-1).
GN ADCYAP1R1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96256640.
RA Hashimoto H., Yamamoto K., Hagigara N., Ogawa N., Nishino A.,
RA Aino H., Nogi H., Imanishi K., Matsuda T., Baba A.;
RT "cDNA cloning of a mouse pituitary adenylate cyclase-activating
RT polypeptide receptor.";
RL Biochim. Biophys. Acta 1281:129-133(1996).
CC -!- FUNCTION: THIS IS A RECEPTOR FOR PACAP-27 AND PACAP-38. THE
CC ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLATE CYCLASE. MAY REGULATE THE RELEASE OF ADRENOCORTICOTROPIN,
CC LUTEINIZING HORMONE, GROWTH HORMONE, PROLACTIN, EPINEPHRINE, AND
CC CATECHOLAMINE. MAY PLAY A ROLE IN SPERMATOGENESIS AND SPERM
CC MOTILITY. CAUSES SMOOTH MUSCLE RELAXATION AND SECRETION IN THE
CC GASTROINTESTINAL TRACT.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D82935; BAA11639.1; -.
CC GCRDB; GCR_1152; -.
CC MGD; MGI:108449; ADCYAP1R1.
CC INTERPRO; IPR000832; -.
CC INTERPRO; IPR002285; -.
CC PFAM; PF00002; 7tm_2; 1.
CC PRINTS; PR00249; GPCRSECRETIN.
CC PRINTS; PR01156; PACAPRECEPTR.
CC PROSITE; PS00649; G_PROTEIN_RECEPT_F2_1; 1.
CC PROSITE; PS00650; G_PROTEIN_RECEPT_F2_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
CC Spermatogenesis.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 468 PITUITARY ADENYLATE CYCLASE ACTIVATING
FT POLYPEPTIDE TYPE I RECEPTOR.
FT DOMAIN 21 155 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 156 178 1 (POTENTIAL).
FT DOMAIN 179 186 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 187 205 2 (POTENTIAL).
FT DOMAIN 206 227 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 228 253 3 (POTENTIAL).
FT DOMAIN 254 268 CYTOPLASMIC (POTENTIAL).

```

RP SEQUENCE FROM N.A.
RX MEDLINE; 94292182.
RA Pausova Z., Bourdon J., Clayton D., Mattei M.-G., Seldin M.F.,
R Janicic N., Riviere M., Szpirer J., Levan G., Szpirer C.;
RT "Cloning of a parathyroid hormone/parathyroid hormone-related peptide
RT receptor (PTHr) cDNA from a rat osteosarcoma (UMR 106) cell line:
RT chromosomal assignment of the gene in the human, mouse, and rat
RT genomes".
RL Genomics 20:20-26(1994).
CC -1- FUNCTION: THIS IS A RECEPTOR FOR PARATHYROID HORMONE AND FOR
CC PARATHYROID HORMONE-RELATED PEPTIDE. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL
CC CYCLASE AND ALSO A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER
CC SYSTEM.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M77184; AAA1811.1; -;
DR EMBL; L19475; AAA68098.1; -;
DR GCRDB; GCR_0206; -;
DR GCRDB; GCR_0938; -;
DR INTERPRO: IPR000832; -;
DR INTERPRO: IPR002170; -;
DR PFAM: PF00002; 7tm_2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PRINTS; PR00393; PTHRMONOMER.
DR PROSITE; PS00649; G-PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G-PROTEIN_RECEP_F2_2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
KW SIGNAL 26 POTENTIAL.
FT CHAIN 1 27 591
FT PARATHYROID HORMONE/PARATHYROID HORMONE-
FT RELATED PEPTIDE RECEPTOR.
FT EXTRACELLULAR (POTENTIAL).
FT 1 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT 2 (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT 3 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT 4 (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT 5 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT 6 (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT 7 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151
FT CARBOHYD 161 161
FT CARBOHYD 166 166
FT CARBOHYD 176 176
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 591 AA; 66260 MW; 21944F3051B9E9C1 CRC64;

Query Match 82.7%; Score 67; DB 1; Length 591;
Best Local Similarity 71.4%; Pred. No. 0.0006;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCNGEVAEKKW 14
:|||||::|
Db 461 FCNGEVAEIRKSW 474

RESULT 8
VIPR_CARAU

ID VIPR_CARAU STANDARD; PRT; 447 AA.
AC Q90308;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE VASOACTIVE INTESTINAL POLYPEPTIDE RECEPTOR (VIP-R) (VIP RECEPTOR).
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprininae; Carassius.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97190233.
RA Chow B.K.C., Yuen T.T.H., Chan K.W.;
RT "Molecular evolution of vertebrate VIP receptors and functional
RT characterization of a VIP receptor from goldfish Carassius auratus";
RL Gen. Comp. Endocrinol. 105:176-185(1997).
CC -1- FUNCTION: THIS IS A RECEPTOR FOR VIP. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL
CC CYCLASE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U56391; AAB05459.1; -;
DR GCRDB; GCR_1205; -;
DR INTERPRO: IPR000832; -;
DR PFAM: PF00002; 7tm_2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PROSITE; PS00649; G-PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G-PROTEIN_RECEP_F2_2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein.
KW DOMA1N 1 103 EXTRACELLULAR (POTENTIAL).
FT TRANSME 104 128
FT 1 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT 2 (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT 3 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT 4 (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT 5 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT 6 (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT 7 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 17 17
FT CARBOHYD 22 22
FT CARBOHYD 64 64
FT CARBOHYD 91 91
FT CARBOHYD 169 169
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 447 AA; 50959 MW; 56839E243702534C CRC64;

Query Match 74.1%; Score 60; DB 1; Length 447;
Best Local Similarity 68.8%; Pred. No. 0.0071;
Matches 11; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 1 YC--NGEVAEKKW 14
|||::|:
Db 352 YCFNGEVAEIRKSW 367

RESULT 9
PACR_HUMAN

```

FT VARIANT 410 410 /FTId-VAR_003582.
FT T -> P (IN MURK JANSEN; CONSISTITUTIVELY
FT ACTIVATED).
FT /FTId-VAR_003583.
FT K -> N (IN REF. 2).
FT CONFLICT 471 471
FT CONFLICT 473 473 S -> C (IN REF. 2).
SO SEQUENCE 593 AA; 66360 MW; DA1400640A6C7F2B CRC64;

Query Match 86.4%; Score 70; DB 1; Length 593;
Best Local Similarity 78.6%; Pred. No. 0.00019; Indels 0; Gaps 0;
Matches 1; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YCNGEVQAEVKKMW 14
Db 461 FCNGEVQAEIRKSW 474

RESULT 6
PTRR_MOUSE
AC P41593; 062119; STANDARD; PRT; 591 AA.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR
DE PRECURSOR (PTH/PTHR RECEPTOR).
GN PTHR1 OR PTHR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-C3H/HEHA.
RX MEDLINE; 95034305.
RA Karperien M., van Dijk T.B., Hoeijmakers T., Cremers F.,
RA Abou-Samra A.-B., Boonstra J., de Laat S.W., Defize L.H.K.;
RT "Expression pattern of parathyroid hormone/parathyroid hormone
RT related peptide receptor mRNA in mouse postimplantation embryos
RT indicates involvement in multiple developmental processes.";
RL Mech. Dev. 47:29-42(1994).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN-BALB/C.
RC STRAIN-BALB/C.
RX MEDLINE; 94255468.
RA McCuaig K.A., Clarke J.C., White J.H.;
RT "Molecular cloning of the gene encoding the mouse parathyroid
RT hormone/parathyroid hormone-related peptide receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:5051-5055(1994).
CC -!- FUNCTION: THIS IS A RECEPTOR FOR PARATHYROID HORMONE AND FOR
CC PARATHYROID HORMONE-RELATED PEPTIDE. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLIL
CC CYCLASE AND ALSO A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER
CC SYSTEM.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X78936; CAA55536.1; -.
CC EMBL; L34611; AAA40011.1; -.
CC EMBL; L34608; AAA40011.1; JOINED.
CC EMBL; L34607; AAA40011.1; JOINED.
CC EMBL; L34609; AAA40011.1; JOINED.
CC EMBL; L34610; AAA40011.1; JOINED.
CC GCRDB; GCR_1005; -.
CC GCRDB; GCR_1614; -.

```

```

DR MGD; MGI:97801; PTHR.
DR INTERPRO; IPR000832; -.
DR PFAM; PF0002; 7tm.2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PRINTS; PR00393; PTHRORMONER.
DR PROSITE; PS00649; G_PROTEIN_RECPE_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECPE_F2_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 26
FT CHAIN 27 591
PTRR_MOUSE
AC P41593; 062119; STANDARD; PRT; 591 AA.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR
DE PRECURSOR (PTH/PTHR RECEPTOR).
GN PTHR1 OR PTHR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1]
RN SEQUENCE FROM N.A.
RC TISSUE-BONE.
RX MEDLINE; 92212903.
RA Abou-Samra A.-B., Jueppner H., Force T., Freeman M.W., Kong X.-F.,
RA Schipani E., Urena P., Richards J., Bonventre J.V., Potts J.T. Jr.,
RA Kronenberg H.M., Segre G.V.;
RT "Expression cloning of a common receptor for parathyroid hormone and
RT parathyroid hormone-related peptide from rat osteoblast-like cells: a
RT single receptor stimulates intracellular accumulation of both cAMP
RT and inositol trisphosphates and increases intracellular free
RT calcium.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2732-2736(1992).
[2]
RN

Query Match 82.7%; Score 67; DB 1; Length 591;
Best Local Similarity 71.4%; Pred. No. 0.0006;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YCNGEVQAEVKKMW 14
Db 461 FCNGEVQAEIRKSW 474

RESULT 7
PTRR_MOUSE
AC P25961; STANDARD; PRT; 591 AA.
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR
DE PRECURSOR (PTH/PTHR RECEPTOR).
GN PTHR1 OR PTHR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1]
RN SEQUENCE FROM N.A.
RC TISSUE-BONE.
RX MEDLINE; 92212903.
RA Abou-Samra A.-B., Jueppner H., Force T., Freeman M.W., Kong X.-F.,
RA Schipani E., Urena P., Richards J., Bonventre J.V., Potts J.T. Jr.,
RA Kronenberg H.M., Segre G.V.;
RT "Expression cloning of a common receptor for parathyroid hormone and
RT parathyroid hormone-related peptide from rat osteoblast-like cells: a
RT single receptor stimulates intracellular accumulation of both cAMP
RT and inositol trisphosphates and increases intracellular free
RT calcium.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2732-2736(1992).
[2]
RN

```


RN RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RX MEDLINE; 93238641.
 RA Schipani E., Karga H., Karaplis A.C., Potts J.T. Jr., Kronenberg H.M.,
 RA Abou-Samra A.-B., Segre G.V., Juppner H.;
 RA "Identical complementary deoxyribonucleic acids encode a human renal
 RT and bone parathyroid hormone (PTH)/PTH-related peptide receptor.";
 RL Endocrinology 132:2157-2165(1993).
 [2]
 RN RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RX MEDLINE; 93367403.
 RA Schneider H., Feyen J.-H., Rao Movva N.;
 RT "Cloning and functional expression of a human parathyroid hormone
 RL receptor.";
 RL Eur. J. Pharmacol. 246:149-155(1993).
 [3]
 RN RP SEQUENCE FROM N.A.
 RC MEDLINE; 95263723.
 RA Schipani E., Weinstein L.S., Bergwitz C., Iida-Klein A., Kong X.F.,
 RA Stuhmann M., Kruse K., Whyte M.P., Murray T., Schmidtkne J., Dop C.,
 RA Brickman A.S., Crawford J.D., Potts J.T. Jr., Kronenberg H.M.,
 RA Abou-Samra A.-B., Segre G.V., Juppner H.;
 RT "Pseudohypoparathyroidism type 1b is not caused by mutations in the
 RT coding exons of the human parathyroid hormone (PTH)/PTH-related
 RT peptide receptor gene.";
 RL J. Clin. Endocrinol. Metab. 80:1611-1621(1995).
 [4]
 RN RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RA Levine M.A.;
 [5]
 RN RP Submitted (xxx-1995) to the EMBL/GenBank/DBJ databases.
 [5]
 RN RP STRUCTURE BY NMR OF 168-198.
 RX MEDLINE; 98409426.
 RA Pellegrini M., Bisello A., Rosenblatt M., Choev M., Mierke D.F.;
 RT "Binding domain of human parathyroid hormone receptor: from
 RT conformation to function.";
 RL Biochemistry 37:12737-12743(1998).
 [6]
 RN RP VARIANT MURK-JANSEN ARG-223 AND PRO-410.
 RX MEDLINE; 96366745.
 RA Schipani E., Langman C.B., Parfitt A.M., Jensen G.S., Kikuchi S.,
 RA Kooh S.W., Cole W.G., Juppner H.;
 RT "Constitutively activated receptors for parathyroid hormone and
 RT parathyroid hormone-related peptide in Jansen's metaphyseal
 RT chondrodysplasia.";
 RL New Engl. J. Med. 335:708-714(1996).
 [8]
 RN RP MUTAGENESIS OF ARG-223 AND PRO-410.
 RX MEDLINE; 97322091.
 RA Schipani E., Jensen G.S., Pincus J., Nissenson R.A., Gardella T.J.,
 RA Juppner H.;
 RT "Constitutive activation of the cyclic adenosine 3',5'-monophosphate
 RT signaling pathway by parathyroid hormone (PTH)/PTH-related peptide
 RT receptors mutated at the two loci for Jansen's metaphyseal
 RT chondrodysplasia.";
 RL Mol. Endocrinol. 11:851-858(1997).
 CC -1- FUNCTION: THIS IS A RECEPTOR FOR PARATHYROID HORMONE AND FOR
 CC PARATHYROID HORMONE-RELATED PEPTIDE. THE ACTIVITY OF THIS
 CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL
 CC CYCLASE AND ALSO A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER
 CC SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MOST TISSUES. MOST ABUNDANT IN
 CC FT

CC CC KIDNEY, BONE AND LIVER.
 CC -1- DISEASE: DEFECTS IN PTHRI ARE THE CAUSE OF BLOMSTRAND TYPE OF
 CC CHONDRODYSPLASIA; A SEVERE SKELETAL DYSPLASIA.
 CC -1- DISEASE: DEFECTS IN PTHRI ARE THE CAUSE OF MURK-JANSEN TYPE OF
 CC METAPHYSEAL CHONDRODYSPLASIA; A SEVERE SKELETAL DYSPLASIA. IT IS
 CC A RARE FORM OF SHORT-LIMBED DWARFISM ASSOCIATED WITH HYPERCALCEMIA
 CC AND NORMAL OR LOW SERUM CONCENTRATIONS OF THE TWO PARATHYROID
 CC HORMONES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L04308; AAA36525.1; -;
 CC EMBL; X68596; CAA48589.1; -;
 CC EMBL; U22409; AAB60657.1; -;
 CC EMBL; U22401; AAB60657.1; JOINED.
 CC EMBL; U22402; AAB60657.1; JOINED.
 CC EMBL; U22403; AAB60657.1; JOINED.
 CC EMBL; U22404; AAB60657.1; JOINED.
 CC EMBL; U22405; AAB60657.1; JOINED.
 CC EMBL; U22406; AAB60657.1; JOINED.
 CC EMBL; U22407; AAB60657.1; JOINED.
 CC EMBL; U22408; AAB60657.1; JOINED.
 CC EMBL; U17418; AAA56774.1; -;
 CC PIR; S29610; S29610.
 CC PIR; A49191; A49191.
 CC PDB; 1BL1; 30-MAR-99.
 CC GCRDB; GCR_0205; -;
 CC GCRDB; GCR_0647; -;
 CC GCRDB; GCR_1335; -;
 CC GCRDB; GCR_2025; -;
 CC MIM; 168468; -;
 CC MIM; 156400; -;
 CC MIM; 215045; -;
 CC INTERPRO; IPR000832; -;
 CC INTERPRO; IPR002170; -;
 CC PFAM; PF00002; 7tm2; 1.
 CC PRINTS; PR00249; GPCRSECRETIN.
 CC PRINTS; PR00393; PTRHORMONER.
 CC PROSITE; PS00649; G_PROTEIN_RECEPT_F2_1; 1.
 CC PROSITE; PS00650; G_PROTEIN_RECEPT_F2_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 CC Disease mutation; 3D-structure.
 CC SIGNAL 1 26
 CC CHAIN 27 593
 CC PARATHYROID HORMONE/PARATHYROID HORMONE-
 CC RELATED PEPTIDE RECEPTOR.
 CC EXTRACELLULAR (POTENTIAL).
 CC 1 (POTENTIAL).
 CC CYTOPLASMIC (POTENTIAL).
 CC 2 (POTENTIAL).
 CC EXTRACELLULAR (POTENTIAL).
 CC 3 (POTENTIAL).
 CC CYTOPLASMIC (POTENTIAL).
 CC 4 (POTENTIAL).
 CC EXTRACELLULAR (POTENTIAL).
 CC 5 (POTENTIAL).
 CC CYTOPLASMIC (POTENTIAL).
 CC 6 (POTENTIAL).
 CC EXTRACELLULAR (POTENTIAL).
 CC 7 (POTENTIAL).
 CC CYTOPLASMIC (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC H -> R (IN MURK JANSEN; CONSTITUTIVELY
 CC ACTIVATED).
 CC FT

DR	EMBL; U55836; AAC32849.1; -.
DR	GCRDB; GCR_14113; -.
DR	INTERPRO; IPR000832; -.
DR	Pfam; PF00002; 7tm.2; 1.
DR	PRINTS; PR00249; GPCRSECRETIN.
DR	PROSITE; PS00649; G_PROTEIN_RECP_F2_1; 1.
DR	PROSITE; PS00650; G_PROTEIN_RECP_F2_2; FALSE_NEG.
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT	SIGNAL 1 24 POTENTIAL
FT	CHAIN 25 546 PARATHYROID HORMONE RECEPTOR.
FT	DOMAIN 27 143 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 144 167 1 (POTENTIAL).
FT	DOMAIN 168 174 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 175 194 2 (POTENTIAL).
FT	DOMAIN 195 235 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 236 258 3 (POTENTIAL).
FT	DOMAIN 259 273 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 274 295 4 (POTENTIAL).
FT	DOMAIN 296 313 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 314 334 5 (POTENTIAL).
FT	DOMAIN 335 361 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 362 380 6 (POTENTIAL).
FT	DOMAIN 381 391 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 392 414 7 (POTENTIAL).
FT	DOMAIN 415 546 CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 106 106 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 116 116 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE 546 AA; 61800 MW; 2825AE4040313527 CRC64;
QY	1 YCNGEVAQVKKMW 14
Db	412 YCNGEVAQVKKTW 425
RESULT 3	
ID PTNR_DIDMA STANDARD; PRT; 585 AA.	
AC P25107;	
DT 01-NOV-1992 (Rel. 22, Created)	
DT 01-MAY-1995 (Rel. 32, Last sequence update)	
DT 15-JUL-1999 (Rel. 38, Last annotation update)	
DE PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR	
DE PRECURSOR (PTH/PTHR RECEPTOR).	
GN PTHR.	
OS Didelphis marsupialis virginiana (North American opossum).	
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.	
[1]	
RP SEQUENCE FROM N.A.	
RX MEDLINE; 92054592.	
RA Jueppner H., Abou-Samra A.-B., Freeman M., Kong X.-F., Schipani E.,	
RA Richards J., Kolakowski L.F. Jr., Hock J., Potts J.T. Jr.,	
RA Kronenberg H.M., Segre G.V.;	
RT "A G protein-linked receptor for parathyroid hormone and parathyroid	
RT hormone-related peptide";	
RL Science 254:1024-1026(1991).	
CC -! FUNCTION: THIS IS A RECEPTOR FOR PARATHYROID HORMONE AND FOR	
CC PARATHYROID HORMONE-RELATED PEPTIDE. THE ACTIVITY OF THIS	
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYL	
CC CYCLASE AND ALSO A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER	
CC SYSTEM.	
CC -! SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.	
CC -! SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.	
CC -----	
CC This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation	

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 09:03:51 ; Search time 58.45 Seconds
(without alignments)
7.652 Million cell updates/sec

Title: US-09-236-468a-2_COPY_415_428

Perfect score: 81

Sequence: 1 YCNEVEQAEVKMW 14

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 87993

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	550	1	PTR2_HUMAN
2	75	92.6	546	1	PTH2_RAT
3	70	86.4	585	1	PTRR_DIDMA
4	70	86.4	585	1	PTRR_PIG
5	70	86.4	593	1	PTRR_HUMAN
6	67	82.7	591	1	PTRR_MOUSE
7	67	82.7	591	1	PTRR_RAT
8	60	74.1	447	1	VIPR_CARAU
9	60	74.1	468	1	PACR_HUMAN
10	60	74.1	496	1	PACR_MOUSE
11	60	74.1	513	1	PACR_BOVIN
12	60	74.1	523	1	PACR_RAT
13	58	71.6	260	1	VIPR_MELGA
14	55	67.9	440	1	SCRC_HUMAN
15	55	67.9	445	1	SCRC_RABIT
16	55	67.9	449	1	SCRC_RAT
17	55	67.9	457	1	VIPR_HUMAN
18	55	67.9	458	1	VIPR_PIG
19	55	67.9	459	1	VIPR_RAT
20	53	65.4	474	1	CALR_RABIT
21	53	65.4	490	1	CALR_HUMAN
22	50	61.7	478	1	CALR_CAVPO
23	48	59.3	411	1	CALR_HUMAN
24	48	59.3	431	1	CRF2_MOUSE
25	48	59.3	437	1	VIPS_MOUSE
26	48	59.3	437	1	VIPS_RAT
27	48	59.3	438	1	VIPS_HUMAN
28	48	59.3	498	1	P25117_HUMAN
29	48	59.3	515	1	CALR_MOUSE
30	48	59.3	516	1	CALR_RAT
31	46	56.8	462	1	GIPR_MESAU
32	45	55.6	411	1	CRF2_RAT
33	45	55.6	463	1	GLPR_HUMAN

RESULT 1

ID	PTR2_HUMAN	STANDARD;	PRT;	550 AA.
AC	P49190;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	PARATHYROID HORMONE RECEPTOR PRECURSOR (PTH2 RECEPTOR).			
GN	PTH2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.			
RC	TISSUE=BRAIN;			
RX	MEDLINE: 95318121.			
RA	Usdin T.B., Gruber C., Bonner T.I.;			
RT	"Identification and functional expression of a receptor selectively recognizing parathyroid hormone, the PTH2 receptor.";			
RL	J. Biol. Chem. 270:15455-15458(1995).			
RN	[2]			
RP	SEQUENCE OF 26-40 AND 306-550 FROM N.A.			
RX	MEDLINE: 97079671.			
RA	Usdin T.B., Modi W., Bonner T.I.;			
RT	"Assignment of the human PTH2 receptor gene (PTH2) to chromosome 2q33 by fluorescence in situ hybridization.";			
RL	Genomics 37:140-141(1996).			
CC	-1- FUNCTION: THIS IS A SPECIFIC RECEPTOR FOR PARATHYROID HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYL CYCLASE.			
CC	-1- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN BRAIN AND PANCREAS. ALSO EXPRESSED IN THE TESTIS.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
EMBL;	U25128; AAC50157.1; -			
DR	EMBL; U47124; AAA96796.1; -			
DR	EMBL; U47129; AAC50767.1; -			
DR	EMBL; U47125; AAC50767.1; JOINED.			
DR	EMBL; U47126; AAC50767.1; JOINED.			
DR	EMBL; U47127; AAC50767.1; JOINED.			
DR	EMBL; U47128; AAC50767.1; JOINED.			
DR	GCRDB; GCR_2003; -			
DR	MIM; 601469; -			
DR	INTERPRO; IPR000832; -			
DR	PFAM; PF00002; 7tm_2; 1.			
DR	PRINTS; PR00249; GPCRSECRETIN.			
DR	PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.			
DR	PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.			
DR				

34	45	55.6	463	1	GLPR_RAT	P32301	rattus norv
35	45	55.6	489	1	GLPR_MOUSE	O35659	mus musculu
36	44	54.3	413	1	CRF2_XENLA	O42603	xenopus lae
37	44	54.3	466	1	GIPR_HUMAN	P48546	homo sapien
38	44	54.3	476	1	VTDB_RABIT	P53789	oryctolagus
39	44	54.3	477	1	GLR_HUMAN	P47871	homo sapien
40	44	54.3	485	1	GLR_MOUSE	Q61606	mus musculu
41	43	53.1	485	1	GLR_RAT	P30082	rattus norv
42	42	51.9	415	1	CRFR_XENLA	O42602	xenopus lae
43	42	51.9	420	1	CRFR_CHICK	O50812	gallus gall
44	42	51.9	461	1	CGRR_HUMAN	Q16602	homo sapien
45	42	51.9	464	1	CGRR_RAT	Q63118	rattus norv

ALIGNMENTS

THIS PAGE BLANK (USPTO)

```

Db 455 YCFLNGEYQAEIKRKW 470
      || |||||:|:|
RESULT 14
JN0902
pituitary adenylate cyclase activating peptide receptor type I precursor - human
C:Species: Homo sapiens (man)
C>Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Jun-2000
C:Accession: JN0902
R:Ogi, K.; Miyamoto, Y.; Masuda, Y.; Habata, Y.; Hosoya, M.; Ohtaki, T.; Masuo, Y.; Onda
Biochem. Biophys. Res. Commun. 196, 1511-1521, 1993
A:Title: Molecular cloning and functional expression of a cDNA encoding a human pituitar
A:Reference number: JN0902; MUID:94071918
A:Accession: JN0902
A:Molecule type: mRNA
A:Residues: 1-525 <OGI>
A:Cross-references: DBJ:017516; NID:9457562; PIDN:BAA04466.1; PID:9540518
A:Experimental source: pituitary
C:Comment: This protein plays pivotal roles as a neurotransmitter and a neuromodulator,
C:Superfamily: glucagon receptor
C:Keywords: glycoprotein; neurotransmitter; receptor
F:1-77/Domain: signal sequence #status predicted <SIG>
F:78-525/Product: pituitary adenylate cyclase activating peptide receptor type I #status
F:105,117,174,357,400,432/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 74.1%; Score 60; DB 2; Length 525;
Best Local Similarity 68.8%; Pred. No. 0.021;
Matches 11; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 1 YC--NGEYQAEVKKMW 14
      || |||||:|:|
Db 457 YCFLNGEYQAEIKRKW 472

RESULT 15
JC2532
secretin receptor - human
C:Species: Homo sapiens (man)
C>Date: 19-Mar-1997 #sequence_revision 11-Apr-1997 #text_change 05-Nov-1999
C:Accession: JC2532
R:Jiang, S.; Ulrich, C.
Biochem. Biophys. Res. Commun. 207, 883-890, 1995
A:Title: Molecular cloning and functional expression of a human pancreatic secretin rece
A:Reference number: JC2532; MUID:95169147
A:Accession: JC2532
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-440 <JIA>
A:Cross-references: EMBL:U20178; NID:9662795; PIDN:AAC50106.1; PID:9662796
A:Experimental source: pancreas
C:Genetics:
A:Gene: GDB:SCFR
A:Cross-references: GDB:270546; OMIM:182098
A:Map position: 2q14.1-2q14.1
C:Superfamily: glucagon receptor

Query Match 67.9%; Score 55; DB 2; Length 440;
Best Local Similarity 68.8%; Pred. No. 0.12;
Matches 11; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 1 YC--NGEYQAEVKKMW 14
      || |||||:|:|
Db 388 YCFLNGEYQAEVKKW 403

```

Search completed: November 8, 2000, 08:53:30
Job time: 361 sec

F:413-433/Domain: transmembrane #status predicted <TM7>
F:47,59,116,299,342,402/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:365,444/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 74.1% Score 60; DB 2; Length 495;
Best Local Similarity 68.8%; Pred. No. 0.019; Mismatches 2; Indels 1; Gaps 1;
Matches 11; Conservative 2

Qy 1 YC--NGEVOAEVKKMW 14
|| |||||:|:|
Db 427 YCFLNGEVOAEIKRKW 442

RESULT 10
A48204
pituitary adenylate cyclase-activating polypeptide type I receptor precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 05-Nov-1999
C:Accession: A48204
R:Pisegna, J.R.; Wank, S.A.
Proc. Natl. Acad. Sci. U.S.A. 90, 6345-6349, 1993
A:Title: Molecular cloning and functional expression of the pituitary adenylate cyclase-
A:Reference number: A48204; MUID:93317678
A:Accession: A48204
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-495 <PTS>
A:Cross-references: GB:L16680; NID:9347941; PIDN:AAA41792.1; PID:g347942
C:Superfamily: glucagon receptor
C:Keywords: anterior pituitary; G protein-coupled receptor; glycoprotein; hypothalamus;
F:1-191/Domain: signal sequence #status predicted <SIG>
F:20-495/Product: pituitary adenylate cyclase-activating polypeptide type I receptor #st

Query Match 74.1% Score 60; DB 2; Length 495;
Best Local Similarity 68.8%; Pred. No. 0.019; Mismatches 2; Indels 1; Gaps 1;
Matches 11; Conservative 2

Qy 1 YC--NGEVOAEVKKMW 14
|| |||||:|:|
Db 427 YCFLNGEVOAEIKRKW 442

RESULT 11
S39061
pituitary adenyl cyclase activating-peptide receptor form 3 - rat
N:Alternate names: PACAP receptor
C:Species: Rattus norvegicus (Norway rat)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Nov-1999
C:Accession: S39061
R:Spengler, D.; Waerber, C.; Pantaloni, C.; Holsboer, F.; Bockaert, J.; Seeburg, P.H.; Jo
Nature 365, 170-175, 1993
A:Title: Differential signal transduction by five splice variants of the PACAP receptor.
A:Reference number: S36768; MUID:93382505
A:Accession: S39061
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-495 <SPE>
A:Cross-references: EMBL:Z23273; NID:g404210; PIDN:CAA80811.1; PID:g404211
C:Superfamily: glucagon receptor

Query Match 74.1% Score 60; DB 2; Length 495;
Best Local Similarity 68.8%; Pred. No. 0.019; Mismatches 2; Indels 1; Gaps 1;
Matches 11; Conservative 2

Qy 1 YC--NGEVOAEVKKMW 14
|| |||||:|:|
Db 427 YCFLNGEVOAEIKRKW 442

RESULT 12

S47631
pituitary adenylate cyclase-activating polypeptide type I receptor precursor - bovine
N:Alternate names: PACAP receptor
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Jun-2000
C:Accession: S47631; S47632
R:Miyaoto, Y.; Habata, Y.; Ohtaki, T.; Masuda, Y.; Ogi, K.; Onda, H.; Fujino, M.
Biochim. Biophys. Acta 1218, 297-307, 1994
A:Title: Cloning and expression of a complementary DNA encoding the bovine receptor f
A:Reference number: S47631; MUID:94325336
A:Accession: S47631
A:Molecule type: mRNA
A:Residues: 1-513 <MIY>
A:Cross-references: EMBL:D17290; NID:g602765; PIDN:BAA04122.1; PID:g1374682
A:Experimental source: brain
A:Accession: S47632
A:Molecule type: protein
A:Residues: 38-41, 'X', 43-50, 'X', 52-66 <MI2>
A:Experimental source: brain
C:Genetics:
A:Introns: 366/2
C:Function:
A:Description: stimulates both adenylate cyclase and phospholipase C
C:Superfamily: glucagon receptor
C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; phosphopr
F:1-37/Domain: signal sequence #status predicted <SIG>
F:38-513/Product: pituitary adenylate cyclase-activating polypeptide type I receptor
F:38-365, 394-513/Product: pituitary adenylate cyclase-activating polypeptide type I r
F:173-195/Domain: transmembrane #status predicted <TM1>
F:204-222/Domain: transmembrane #status predicted <TM2>
F:245-270/Domain: transmembrane #status predicted <TM3>
F:286-308/Domain: transmembrane #status predicted <TM4>
F:326-349/Domain: transmembrane #status predicted <TM5>
F:396-416/Domain: transmembrane #status predicted <TM6>
F:431-451/Domain: transmembrane #status predicted <TM7>
F:65,77,134,360,420/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:383,462/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 74.1% Score 60; DB 2; Length 513;
Best Local Similarity 68.8%; Pred. No. 0.02; Mismatches 2; Indels 1; Gaps 1;
Matches 11; Conservative 2

Qy 1 YC--NGEVOAEVKKMW 14
|| |||||:|:|
Db 445 YCFLNGEVOAEIKRKW 460

RESULT 13
S39060
pituitary adenyl cyclase activating-peptide receptor form 2 - rat
N:Alternate names: PACAP receptor
C:Species: Rattus norvegicus (Norway rat)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
C:Accession: S39060
R:Spengler, D.; Waerber, C.; Pantaloni, C.; Holsboer, F.; Bockaert, J.; Seeburg, P.H.;
Nature 365, 170-175, 1993
A:Title: Differential signal transduction by five splice variants of the PACAP recept
A:Reference number: S36768; MUID:93382505
A:Accession: S39060
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-523 <SPE>
A:Cross-references: EMBL:Z23272; NID:g404195; PIDN:CAA80810.1; PID:g404196
C:Superfamily: glucagon receptor

Query Match 74.1% Score 60; DB 2; Length 523;
Best Local Similarity 68.8%; Pred. No. 0.02; Mismatches 2; Indels 1; Gaps 1;
Matches 11; Conservative 2

Qy 1 YC--NGEVOAEVKKMW 14

A:Molecule type: mRNA
A:Residues: 1-381 <SVO>
A:Cross-references: EMBL:222735; NID:g311228; PIDN:CAA80429.1; PID:g311229
C:Superfamily: glucagon receptor

Query Match 74.1%; Score 60; DB 2; Length 381;
Best Local Similarity 68.8%; Pred. No. 0.015;
Matches 11; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

OY 1 YC--NGEQVAEVKKMW 14
|| |||||:|:|
Db 313 YCFLNGEQVAEIKRW 328

RESULT 7

JN0616

pituitary adenylate cyclase-activating polypeptide receptor 46-5 precursor - rat

N:Alternate names: PACAP receptor
C:Species: Rattus norvegicus (Norway rat)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 08-Oct-1999
C:Accession: JN0616; S36768
R:Hosoya, M.; Onda, H.; Ogi, K.; Masuda, Y.; Miyamoto, Y.; Ohtaki, T.; Okazaki, H.; Arimura, H.; Lutz, E.M.; West, K.M.; Fink, G.; Harmar, A.J.
Biochem. Biophys. Res. Commun. 194, 133-143, 1993
A:Title: Molecular cloning and functional expression of rat cDNAs encoding the receptor
A:Reference number: JN0616; MUID:93326107
A:Accession: JN0616
A:Molecule type: mRNA
A:Residues: 1-467 <HOS>
A:Experimental source: brain
R:Spengler, D.; Waerber, C.; Pantaloni, C.; Holsboer, F.; Bockaert, J.; Seeburg, P.H.; Jozefowicz, J.F.
Nature 365, 170-175, 1993
A:Title: Differential signal transduction by five splice variants of the PACAP receptor
A:Reference number: S36768; MUID:93382505
A:Accession: S36768

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-467 <SPR>
A:Cross-references: EMBL:223279; NID:g404252; PIDN:CAA80817.1; PID:g404253
C:Superfamily: glucagon receptor
C:Keywords: alternative splicing; glycoprotein; receptor
F:1-19/domain: signal sequence #status predicted <SIG>
F:20-467/product: pituitary adenylate cyclase-activating polypeptide receptor 46-5 #status predicted
F:47-59,116,299,342,374/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 74.1%; Score 60; DB 2; Length 467;
Best Local Similarity 68.8%; Pred. No. 0.018;
Matches 11; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

OY 1 YC--NGEQVAEVKKMW 14
|| |||||:|:|
Db 399 YCFLNGEQVAEIKRW 414

RESULT 8

S39063

pituitary adenylate cyclase activating-peptide receptor form 5 - rat

N:Alternate names: PACAP receptor
C:Species: Rattus norvegicus (Norway rat)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Nov-1999
C:Accession: S39063
R:Spengler, D.; Waerber, C.; Pantaloni, C.; Holsboer, F.; Bockaert, J.; Seeburg, P.H.; Jozefowicz, J.F.
Nature 365, 170-175, 1993
A:Title: Differential signal transduction by five splice variants of the PACAP receptor
A:Reference number: S36768; MUID:93382505
A:Accession: S39063
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-494 <SPR>
A:Cross-references: EMBL:223275; NID:g404197; PIDN:CAA80813.1; PID:g404198
C:Superfamily: glucagon receptor

Query Match 86.4%; Score 70; DB 2; Length 585;
Best Local Similarity 78.6%; Pred. No. 0.00047;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCNGEVOAEVKKMW 14
:|||||:|||||
Db 455 FCNGEVOAEIRKSW 468

RESULT 3

A49191
parathyroid hormone/PTH-related peptide receptor - human
N:Alternate names: parathyroid hormone/parathyroid related peptide receptor
C:Species: Homo sapiens (man)
C>Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 11-Jan-2000
C:Accession: I38139; A49191; I38113; G01562; S29610
R:Schipani, E.; Weinstein, L.S.; Bergwitz, C.; Iida-Klein, A.; Kong, X.F.; Stuhmann, M.; Kronenberg, H.M.; Abou-Samra, A.B.; Segre, G.V.; Jueppner, H.
J. Clin. Endocrinol. Metab. 80, 1611-1621, 1995
A:Title: Pseudohypoparathyroidism type Ib is not caused by mutations in the coding exons
A:Reference number: I38139; MUID:95263723
A:Accession: I38139
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-593 <RES>
A:Cross-references: EMBL:U22409; NID:g987594; PIDN:AA860657.1; PID:g987596
R:Schipani, E.; Karga, H.; Karapilis, A.C.; Potts Jr., J.T.; Kronenberg, H.M.; Segre, G.V.
Endocrinology 132, 2157-2165, 1993
A:Title: Identical complementary deoxyribonucleic acids encode a human renal and bone pa
A:Reference number: A49191; MUID:93238641
A:Accession: A49191
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-593 <SCH>
A:Cross-references: GB:L04308; NID:g190721; PIDN:AAA36525.1; PID:g190722
A:Note: sequence extracted from NCBI backbone (NCBIN:130233, NCBIP:130234)
R:Schneider, H.; Feyen, J.H.; Seuwen, K.; Movva, N.R.
Eur. J. Pharmacol. 246, 149-155, 1993
A:Title: Cloning and functional expression of a human parathyroid hormone receptor.
A:Reference number: I38113; MUID:93387403
A:Accession: I38113
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-593 <RE2>
A:Cross-references: EMBL:X68596; NID:g396812; PIDN:CAA48589.1; PID:g396813
R:Levine, M.
submitted to the EMBL Data Library, November 1994
A:Reference number: G07787
A:Accession: G01562
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-593 <LEV>
A:Cross-references: EMBL:U17418; NID:g596129; PIDN:AAA56774.1; PID:g596130
C:Genetics:
A:Introns: 25/3; 60/1; 105/1; 142/1; 181/3; 213/2; 278/3; 330/1; 350/2; 372/3; 404/2; 45
A:Superfamily: glucagon receptor
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 86.4%; Score 70; DB 2; Length 593;
Best Local Similarity 78.6%; Pred. No. 0.00047;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCNGEVOAEVKKMW 14
:|||||:|||||
Db 461 FCNGEVOAEIRKSW 474

RESULT 4

S44203
parathyroid hormone-related peptide receptor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 11-Jan-2000

C:Accession: S44203
R:Karprien, M.; van Dijk, T.B.; Hoeijmakers, T.; Cremers, F.; Abou-Samra, A.B.; Boon
submitted to the EMBL Data Library, April 1994
A:Description: Expression pattern of parathyroid hormone/parathyroid hormone related
A:Reference number: S44203
A:Accession: S44203
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-591 <RAR>
A:Cross-references: EMBL:X78936; NID:g474828; PIDN:CAA55536.1; PID:g474829
C:Superfamily: glucagon receptor

Query Match 82.7%; Score 67; DB 2; Length 591;
Best Local Similarity 71.4%; Pred. No. 0.0015;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCNGEVOAEVKKMW 14
:|||||:|||||
Db 461 FCNGEVOAEIRKSW 474

RESULT 5

I54195
parathyroid hormone/parathyroid hormone related-peptide receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 11-Jan-2000
C:Accession: I54195; A42698
R:Pausova, Z.; Bourdon, J.; Clayton, D.; Mattei, M.G.; Seldin, M.F.; Janicic, N.; Riv
Genomics 20, 20-26, 1994
A:Title: Cloning of a parathyroid hormone/parathyroid hormone-related peptide recepto
and rat genomes.
A:Reference number: I54195; MUID:94292182
A:Accession: I54195
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-591 <RES>
A:Cross-references: GB:L19475; NID:g467316; PIDN:AAA68098.1; PID:g467317
R:Abou-Samra, A.B.; Juppner, H.; Force, T.; Freeman, M.W.; Kong, X.F.; Schipani, E.;
Proc. Natl. Acad. Sci. U.S.A. 89, 2732-2736, 1992
A:Title: Expression cloning of a common receptor for parathyroid hormone and parathyr
n of both cAMP and inositol trisphosphates and increases intracellular free calcium.
A:Reference number: A42698; MUID:92212903
A:Accession: A42698
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-585, 'G', 587-591 <ABO>
A:Experimental source: ROS 17/2.8 osteosarcoma cells
A:Note: sequence extracted from NCBI backbone (NCBIP:92187)
C:Superfamily: glucagon receptor

Query Match 82.7%; Score 67; DB 2; Length 591;
Best Local Similarity 71.4%; Pred. No. 0.0015;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCNGEVOAEVKKMW 14
:|||||:|||||
Db 461 FCNGEVOAEIRKSW 474

RESULT 6

S33449
pituitary adenylate cyclase-activating polypeptide receptor homolog - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 11-Jan-2000
C:Accession: S33449
R:Svoboda, M.; Ciccarelli, E.; Tastenoy, M.; Christophe, J.
submitted to the EMBL Data Library, May 1993
A:Description: Molecular cloning of a PACAP-type receptor.
A:Reference number: S33449
A:Accession: S33449
A:Status: preliminary

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model
Run on: November 8, 2000, 08:53:30 ; Search time 99.87 Seconds
(without alignments)
8.896 Million cell updates/sec

Title: US-09-236-468a-2_COPY_415_428

Perfect score: 81

Sequence: 1 YCNGEVOAEVKKMW 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 182106

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_65.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	550	A57519	parathyroid hormon
2	70	86.4	585	A39286	parathyroid hormon
3	70	86.4	593	A49191	parathyroid hormon
4	67	82.7	591	S44203	parathyroid hormon
5	67	82.7	591	I54195	parathyroid hormon
6	60	74.1	381	S33449	pituitary adenylat
7	60	74.1	467	JN0616	pituitary adenylat
8	60	74.1	494	S39063	pituitary adenylat
9	60	74.1	495	S36114	pituitary adenylat
10	60	74.1	495	A48204	pituitary adenylat
11	60	74.1	495	S39061	pituitary adenylat
12	60	74.1	513	S47631	pituitary adenylat
13	60	74.1	523	S39060	pituitary adenylat
14	60	74.1	525	JN0902	pituitary adenylat
15	55	67.9	440	JC2532	secretin receptor
16	55	67.9	449	S16319	secretin receptor
17	55	67.9	459	JH0594	vasoactive intesti
18	55	67.9	460	JC2194	vasoactive intesti
19	55	67.9	495	JC2195	vasoactive intesti
20	53	65.4	474	I37317	calcitonin recepto
21	53	65.4	490	S34866	calcitonin recepto
22	48	59.3	431	I49279	sausage/corticot
23	48	59.3	431	I49149	CRF receptor - mou
24	48	59.3	437	JU0185	PACAP/VIP receptor
25	48	59.3	437	S39069	vasoactive intesti
26	48	59.3	438	G02822	vasoactive intesti
27	48	59.3	478	A37430	calcitonin recepto
28	48	59.3	479	S33746	calcitonin recepto
29	48	59.3	482	A39285	calcitonin recepto

30	48	59.3	498	2	I47130	calcitonin recepto
31	48	59.3	515	2	I60800	calcitonin recepto
32	48	59.3	515	2	I49154	calcitonin recepto
33	48	59.3	589	2	I59297	parathyroid hormon
34	46	56.8	462	2	JC2462	corticotropic-in-rele
35	45	55.6	411	2	A55610	glucagon-like pept
36	45	55.6	463	2	S71624	glucagon-like pept
37	45	55.6	463	2	A46172	glucagon-like pept
38	45	55.6	463	2	JN0807	glucagon-like pept
39	45	55.6	463	2	I84494	glucagon-like pept
40	44	54.3	466	2	G02234	gastric inhibitory
41	44	54.3	466	2	S66676	glucose-dependent
42	44	54.3	477	2	JC2041	glucagon receptor
43	44	54.3	485	2	JC4363	glucagon receptor
44	44	54.3	764	2	T05409	hypothetical prote
45	43	53.1	485	2	JQ1957	glucagon receptor

ALIGNMENTS

RESULT 1

A57519 parathyroid hormone receptor 2 precursor - human

N:Alternate names: PTH2 receptor

C:Species: Homo sapiens (man)

C>Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 24-Nov-1999

C:Accession: A57519

R:Usdin, T.B.; Gruber, C.; Bonner, T.I.

J. Biol. Chem. 270, 15455-15458, 1995

A:Title: Identification and functional expression of a receptor selectively recognizi

A:Reference number: A57519; MUID:95318121

A:Accession: A57519

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-550 <USD>

A:Cross-references: GB:025128; NID:g887966; PIDN:AAC50157.1; PID:g887967

C:Genetics:

A:Gene: GDB:PTH2; PTHR2R

A:Cross-references: GDB:731977; OMIM:601469

A:Map position: 2q33-2q33

C:Superfamily: glucagon receptor

C:Keywords: hormone receptor

Query Match 100.0%; Score 81; DB 2; Length 550;
Best Local Similarity 100.0%; Pred. No. 6.1e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YCNGEVOAEVKKMW 14

Db 415 YCNGEVOAEVKKMW 428

RESULT 2

A39286

parathyroid hormone / parathyroid hormone-related peptide - North American opossum

C:Species: Didelphis virginiana, Didelphis marsupialis virginiana (North American opo

C>Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 11-Jan-2000

C:Accession: A39286

R:Jueppner, H.; Abou-Samra, A.B.; Freeman, M.; Kong, X.F.; Schipani, E.; Richards, J.

Science 254, 1024-1026, 1991

A:Title: A G protein-linked receptor for parathyroid hormone and parathyroid hormone-

A:Reference number: A39286; MUID:92054592

A:Accession: A39286

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-585 <JUE>

A:Cross-references: GB:M74445

C:Superfamily: glucagon receptor

C:Keywords: G protein-coupled receptor; transmembrane protein

THIS PAGE BLANK (USPTO)

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY F
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 994-1700
TELEFAX: (201) 994-1744
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-465-976A-3

Query Match 74.1%; Score 60; DB 2; Length 376;
Best Local Similarity 68.8%; Pred. No. 0.013;
Matches 11; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

Qy 1 YC--NGEVOAEVKKMW 14
|| |||||:|:|
Db 307 YCFLNGEVOAEIKRKW 322

RESULT 15

US-08-982-412-3
Sequence 3, Application US/08982412
Patent No. 5958729
GENERAL INFORMATION:
APPLICANT: SOPPET, DANIEL R
APPLICANT: LI, YI
APPLICANT: ROSEN, CRAIG A
APPLICANT: RUBEN, STEVEN M
TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE,
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/982,412
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF181PCT2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-982-412-3

Query Match 74.1%; Score 60; DB 2; Length 376;
Best Local Similarity 68.8%; Pred. No. 0.013;

Matches 11; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
Qy 1 YC--NGEVOAEVKKMW 14
|| |||||:|:|
Db 307 YCFLNGEVOAEIKRKW 322

Search completed: November 8, 2000, 08:49:19
Job time: 112 sec

APPLICANT: KITADA, Chieko
APPLICANT: ISHIBASHI, Yoshihiro
APPLICANT: HOSOYA, Masaki
APPLICANT: OGI, Kazuhiro
APPLICANT: MIYAMOTO, Yasunori
APPLICANT: HABATA, Yugo
APPLICANT: SHIMAMOTO, No. 585878710
TITLE OF INVENTION: PACAP RECEPTOR PROTEIN, METHOD FOR PREPARING
TITLE OF INVENTION: SAID PROTEIN, AND USE THEREOF
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
ADDRESSEE: CUSHMAN
STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,897A
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/202,986
FILING DATE: February 25, 1994
ATTORNEY/AGENT INFORMATION:
NAME: RESNICK, David S.
REGISTRATION NUMBER: 34235
REFERENCE/DOCKET NUMBER: 44168-DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-811-897A-11

Query Match 74.1%; Score 60; DB 2; Length 76;
Best Local Similarity 68.8%; Pred. No. 0.0024;
Matches 11; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 1 YC--NGEVQAEVKRW 14
|| |||||:|
Db 52 YCFLNGEVQAEIKRW 67

RESULT 13
US-08-855-213-11
Sequence 11, Application US/08855213
Patent No. 5892004
GENERAL INFORMATION:
APPLICANT: ONDA, Haruo
APPLICANT: OHTAKI, Tetsuya
APPLICANT: MASUDA, Yasushi
APPLICANT: KITADA, Chieko
APPLICANT: ISHIBASHI, Yoshihiro
APPLICANT: HOSOYA, Masaki
APPLICANT: OGI, Kazuhiro
APPLICANT: MIYAMOTO, Yasunori
APPLICANT: HABATA, Yugo
APPLICANT: SHIMAMOTO, No. 589200410
TITLE OF INVENTION: PACAP RECEPTOR PROTEIN, METHOD FOR
PREPARING SAID PROTEIN, AND USE THEREOF
NUMBER OF SEQUENCES: 55

CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
ADDRESSEE: CUSHMAN
STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/855,213
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/202,986
FILING DATE: 25-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: RESNICK, David S.
REGISTRATION NUMBER: 34235
REFERENCE/DOCKET NUMBER: 44168
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-855-213-11

Query Match 74.1%; Score 60; DB 2; Length 76;
Best Local Similarity 68.8%; Pred. No. 0.0024;
Matches 11; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 1 YC--NGEVQAEVKRW 14
|| |||||:|
Db 52 YCFLNGEVQAEIKRW 67

RESULT 14
US-08-465-976A-3
Sequence 3, Application US/08465976A
Patent No. 5869632
GENERAL INFORMATION:
APPLICANT: SOPPET, DANIEL R
APPLICANT: LI, YI
APPLICANT: ROSEN, CRAIG A
APPLICANT: RUBEN, STEVEN M
TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN GILFILLAN, CECCHI
ADDRESSEE: STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NJ
COUNTRY: US
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,976A
FILING DATE: 06-JUN-1995

; FILING DATE: 04-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/071003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-8906
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-249A-19

Query Match 86.4%; Score 70; DB 2; Length 585;
Best Local Similarity 78.6%; Pred. No. 0.00043;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YCNGEVQAEVKKMW 14
:|||||:|
Db 455 FCNGEVQAEIRKSW 468

RESULT 10
US-08-468-249A-21
; Sequence 21, Application US/08468249A
; Patent No. 5886148
; GENERAL INFORMATION:
; APPLICANT: Segre et al., Gino V.
; TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR AND DNA
; TITLE OF INVENTION: ENCODING SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,249A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,475
; FILING DATE: 06-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,702
; FILING DATE: 04-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/071003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-8906
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-249A-21

Query Match 86.4%; Score 70; DB 2; Length 593;
Best Local Similarity 78.6%; Pred. No. 0.00044;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YCNGEVQAEVKKMW 14
:|||||:|
Db 461 FCNGEVQAEIRKSW 474

RESULT 11
US-08-468-249A-20
; Sequence 20, Application US/08468249A
; Patent No. 5886148
; GENERAL INFORMATION:
; APPLICANT: Segre et al., Gino V.
; TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR AND DNA
; TITLE OF INVENTION: ENCODING SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,249A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,475
; FILING DATE: 06-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,702
; FILING DATE: 04-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/071003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-249A-20

Query Match 82.7%; Score 67; DB 2; Length 591;
Best Local Similarity 71.4%; Pred. No. 0.0014;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YCNGEVQAEVKKMW 14
:|||||:|
Db 461 FCNGEVQAEIRKSW 474

RESULT 12
US-08-811-897A-11
; Sequence 11, Application US/08811897A
; Patent No. 5858787
; GENERAL INFORMATION:
; APPLICANT: ONDA, Haruo
; APPLICANT: OHTAKI, Tetsuya
; APPLICANT: MASUDA, Yasushi

ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: 699 Prince Street
CITY: Alexandria
STATE: Virginia
COUNTRY: US
ZIP: 22313
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142,551B
FILING DATE: 25-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/077,296
FILING DATE: 14-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/898,219
FILING DATE: 12-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,677
FILING DATE: 22-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Swiss, Gerald F.
REGISTRATION NUMBER: 30,113
REFERENCE/DOCKET NUMBER: 000324-010
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 854-7400
TELEFAX: (415) 854-8275
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..585
OTHER INFORMATION: /note= "PTH receptor"
US-08-142-551B-125

Query Match 86.4%; Score 70; DB 2; Length 585;
Best Local Similarity 78.6%; Pred. No. 0.00043;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YCNGEVQAEVKKMW 14
:|||||:|
Db 455 FCNGEVQAEIKKSW 468

RESULT 8
US-08-869-477-6
Sequence 6, Application US/08869477
Patent No. 5846747
GENERAL INFORMATION:
APPLICANT: Thorens, Bernard
TITLE OF INVENTION: Receptor for the Glucagon-Like-Peptide-1
TITLE OF INVENTION: (GLP-1)
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5846747o No. 5846747disk of No. 5846747th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/869,477
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/142,439
FILING DATE: 24-NOV-93
APPLICATION NUMBER: DK 398/92
FILING DATE: 25-MAR-92
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP93/00697
FILING DATE: 23-MAR-93
ATTORNEY/AGENT INFORMATION:
NAME: Harrington, James J.
REGISTRATION NUMBER: 38,711
REFERENCE/DOCKET NUMBER: 3756.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Didelphis virginiana
US-08-869-477-6

Query Match 86.4%; Score 70; DB 2; Length 585;
Best Local Similarity 78.6%; Pred. No. 0.00043;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YCNGEVQAEVKKMW 14
:|||||:|
Db 455 FCNGEVQAEIKKSW 468

RESULT 9
US-08-468-249A-19
Sequence 19, Application US/08468249A
Patent No. 5886148
GENERAL INFORMATION:
APPLICANT: Segre et al., Gino V.
TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR AND DNA
TITLE OF INVENTION: ENCODING SAME
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,249A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,475
FILING DATE: 06-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/681,702

; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 59 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-011A-24

Query Match 86.4%; Score 70; DB 3; Length 59;
Best Local Similarity 78.6%; Pred. No. 3.9e-05;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YCNGEVOAEVKKMW 14
:|||||:|
Db 29 FCNGEVOAEIKKSW 42

RESULT 5
US-08-468-249A-18
; Sequence 18, Application US/08468249A
; Patent No. 5886148
; GENERAL INFORMATION:
; APPLICANT: Segre et al., Gino V.
; TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR AND DNA
; TITLE OF INVENTION: ENCODING SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,249A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,475
; FILING DATE: 06-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,702
; FILING DATE: 04-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/071003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 515 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-249A-18

Query Match 86.4%; Score 70; DB 2; Length 515;
Best Local Similarity 78.6%; Pred. No. 0.00038;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YCNGEVOAEVKKMW 14
:|||||:|
Db 455 FCNGEVOAEIKKSW 468

RESULT 6
US-08-142-439A-6
; Sequence 6, Application US/08142439A
; Patent No. 5670360
; GENERAL INFORMATION:
; APPLICANT: Thorens, Bernard
; TITLE OF INVENTION: Receptor for the Glucagon-Like-Peptide-1
; TITLE OF INVENTION: (GLP-1)
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5670360o No. 5670360disk of No. 5670360th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,439A
; FILING DATE: 24-NOV-93
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 398/92
; FILING DATE: 25-MAR-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/00697
; FILING DATE: 23-MAR-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Harrington, James J.
; REGISTRATION NUMBER: 38,711
; REFERENCE/DOCKET NUMBER: 3756,204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Didelphis virginiana
US-08-142-439A-6

Query Match 86.4%; Score 70; DB 1; Length 585;
Best Local Similarity 78.6%; Pred. No. 0.00043;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YCNGEVOAEVKKMW 14
:|||||:|
Db 455 FCNGEVOAEIKKSW 468

RESULT 7
US-08-142-551B-125
; Sequence 125, Application US/08142551B
; Patent No. 5814603
; GENERAL INFORMATION:
; APPLICANT: Oldenburg, Kevin R.
; APPLICANT: Selick, Harold E.
; TITLE OF INVENTION: COMPOUNDS WITH PTH ACTIVITY AND
; TITLE OF INVENTION: RECOMBINANT DNA VECTORS ENCODING SAME
; NUMBER OF SEQUENCES: 132
; CORRESPONDENCE ADDRESS:


```
Db 29 YCNGEVOAEVKKMW 42
;
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07085
; FILING DATE: 05-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-393
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-07085-2

Query Match 100.0%; Score 81; DB 4; Length 541;
Best Local Similarity 100.0%; Pred. No. 5.7e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YCNGEVOAEVKKMW 14
| | | | | | | | | |
Db 415 YCNGEVOAEVKKMW 428

RESULT 4
US-08-468-011A-24
; Sequence 24, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-011A-2

Query Match 100.0%; Score 81; DB 3; Length 541;
Best Local Similarity 100.0%; Pred. No. 5.7e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YCNGEVOAEVKKMW 14
| | | | | | | | | |
Db 415 YCNGEVOAEVKKMW 428

RESULT 3
PCT-US95-07085-2
; Sequence 2, Application PC/TUS9507085
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone Receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
```

QY 1 YCNGEVQAEVKKMW 14

Search completed: November 8, 2000, 08:51:46
Job time: 258 sec

Db 52 ycfIngevqaeikrk 67
|| |||||:|:|
RESULT 15
W59667
ID W59667 standard; Protein; 376 AA.
XX AC W59667;
XX 12-OCT-1998 (first entry)
XX DE Amino acid sequence of rat PACAP-like receptor.
XX Rat; PACAP-like receptor; G-protein coupled receptor; HCEGH45;
KW hypersecretory condition; pituitary adenylate cyclase; agonist;
KW memory; amnesia; nerve cell death; neuropathy; Alzheimer's disease;
KW antagonist.
XX OS Mus sp.
XX WO9824900-A1.
XX PN 11-JUN-1998.
XX PD 21-NOV-1997; 97WO-US20547.
XX PF 02-DEC-1996; 96US-0032186.
XX PR (HUMA-) HUMAN GENOME SCI INC.
XX PA Li Y, Rosen CA, Ruben SM, Soppet DR;
XX PI WPI; 1998-333320/29.
XX DR New isolated human G-protein receptor, HCEGH45 - used to develop
XX products for treating e.g. hypersecretory conditions, to improve
XX memory, to treat amnesia or to prevent nerve cell death in
XX neuropathy
XX PS Disclosure; Fig 3; 82pp; English.
XX This is the amino acid sequence of the rat pituitary adenylate
CC cyclase activating polypeptide (PACAP) used in the method of the
CC invention to develop products to treat hypersecretory conditions.
CC The novel protein is a human G-protein pituitary adenylate cyclase
CC activating polypeptide (PACAP)-like receptor for amnesia like
CC neuropeptides, designated HCEGH45. Agonists can be used to treat
CC hypersecretory conditions, to improve memory, to treat amnesia and
CC prevent nerve cell death in neuropathy to prevent and/or treat
CC diseases such as Alzheimer's disease. G-protein coupled receptor
CC agonists can also be used to treat asthma, Parkinson's disease, acute
CC heart failure, hypotension, urinary retention, and osteoporosis.
CC Antagonists can be used to treat hypersecretory conditions and to
CC create pharmacological amnesia or effect long-term memory. G-protein
CC coupled receptor antagonists can also be used to treat hypertension,
CC angina pectoris, myocardial infarction, ulcers, asthma, allergies,
CC psychoses, depression, migraine, vomiting, and benign prostatic
CC hypertrophy. The products can also be used for detection, diagnosis
CC and drug screening.
XX SQ Sequence 376 AA;

Query Match 74.1%; Score 60; DB 19; Length 376;
Best Local Similarity 68.8%; Pred No. 0.016; 1; Indels 2; Gaps 1;
Matches 11; Conservative 2; Mismatches 1;
Qy 1 YC--NGEVQAEVKKRW 14
|| |||||:|:|
Db 307 ycfIngevqaeikrk 322

Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YCNGEVAEVRKKMW 14
:|||||::|
Db 461 fcngvqaeirksw 474

RESULT 13

R27707
ID R27707 standard; Protein; 614 AA.

XX XX

AC R27707;

XX XX

DT 16-MAR-1993 (first entry)

XX XX

DE Human kidney PTH/PTHrP receptor.

XX XX

KW Parathyroid hormone; related protein; calcium; antagonist;
KW antibodies; hypercalcaemia.

XX XX

OS Homo sapiens.

XX XX

PN WO9217602-A.

XX XX

PD 15-OCT-1992.

XX XX

PF 06-APR-1992; 92WO-US02821.

XX XX

PR 05-APR-1991; 91US-0681702.

XX XX

PR 06-APR-1992; 92US-0864475.

XX XX

PA (GENO) GEN HOSPITAL CORP OFFICE TECHNOLOGY.

XX XX

PI Abou-samra A, Juppner H, Kronenberg HM, Potts JT, Schipani E;
PI Segre GV;

XX XX

DR WPI; 1992-366271/44.

XX XX

DR N-PSDB; Q29607.

XX XX

PT New DNA encoding parathyroid hormone receptor, DNA and antibodies
PT - for (differential) diagnosis of hypercalcaemia, and diagnosis
PT and treatment of tumours

XX XX

PS Claim 22; Fig 6; 91pp; English.

XX XX

CC The protein sequence was deduced from the cDNA sequence obtd by
CC screening a human kidney oligo dt-primed cDNA library in lambda
CC gt10 and a genomic library of human placental DNA in EMBL3 with a
CC probe comprising most of the coding sequence of rat bone parathyroid
CC hormone/parathyroid hormone related protein (PTH/PTHrP) receptor
CC protein. The clone encodes a protein which may be used in a
CC therapeutic compsn. to inhibit activation of PTH or PTHrP and thus
CC reduce the level of calcium in the blood. Cpsd. capable of competing
CC with PTH or PTHrP for binding can be identified using the protein prod.
CC and DNAs homologous to PTH DNA can be identified using fragments of the
CC clone as probes. The sequence may be used for the prodn. of antibodies
CC useful for the treatment, classification, prognosis and/or treatment of
CC disorders related to the interaction between a cell receptor and a
CC ligand such as in hypercalcaemia. See also R27704-16.

XX XX

SQ Sequence 614 AA;

Query Match 77.8%; Score 63; DB 13; Length 614;
Best Local Similarity 75.0%; Pred. No. 0.008;
Matches 12; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

Qy 1 YC--NGEVAEVRKKMW 14
|| |||||::|
Db 459 ycfngvqaeirksw 474

RESULT 14

R58653

ID R58653 standard; Peptide; 76 AA.

XX XX

AC R58653;

XX XX

DT 12-MAY-1995 (first entry)

XX XX

DE PACAP receptor protein homology region 11.

XX XX

KW Bovine; pituitary adenylate cyclase activating polypeptide; PACAP;

XX XX

KW adenylate cyclase; receptor; type 1A; signal sequence;

XX XX

KW hydrophobic cluster; transmembrane; human; PACAP receptor; pituitary;

XX XX

KW gland-derived; diagnosis; neuropathy; Alzheimer's disease; gene therapy;

XX XX

OS Synthetic.

XX XX

PN EP618291-A.

XX XX

PD 05-OCT-1994.

XX XX

PF 24-FEB-1994; 94EP-0102757.

XX XX

PR 26-FEB-1993; 93JP-0038755.

XX XX

PR 05-APR-1993; 93JP-0078290.

XX XX

PR 27-APR-1993; 93JP-0100669.

XX XX

PR 17-MAY-1993; 93JP-0114446.

XX XX

PR 24-JUN-1993; 93JP-0153963.

XX XX

PR 10-NOV-1993; 93JP-0281413.

XX XX

PR 27-DEC-1993; 93JP-0333175.

XX XX

PA (TAKE) TAKEDA CHEM IND LTD.

XX XX

PI Habata Y, Hosoya M, Ishibashi Y, Kitada C, Masuda Y;

XX XX

PI Miyamoto Y, Ogi K, Ohtaki T, Onda H, Shimamoto N;

XX XX

DR WPI; 1994-304460/38.

XX XX

PT New pituitary adenylate cyclase activating polypeptide receptor

XX XX

PT proteins - used to develop prods. for use in the diagnosis and

XX XX

PT treatment of neuropathy such as Alzheimer's disease

XX XX

PS Claim 2; Page 51; 164pp; English.

XX XX

CC A series of peptide sequences (R58643-54) describing the regions of high
CC homology between the human, bovine and rat pituitary adenylate cyclase
CC activating polypeptide (PACAP) receptor proteins. The genes (Q72203-10)
CC encode proteins (R58655-62) of around 500-550 amino acids. The proteins
CC contain a signal sequence which may be cleaved to produce mature proteins
CC (R58663-71). Hydrophobic measurements have deduced that the proteins
CC contain 7 hydrophobic clusters considered to be transmembrane domains in
CC tandem. The gene for the bovine PACAP receptor was cloned by purifying
CC the protein to homogeny. A probe was constructed to match the N-terminal
CC region which was subsequently used to probe a bovine brain cDNA library
CC to obtain a cDNA clone of the PACAP receptor. The DNA sequence was used
CC to obtain human PACAP receptor cDNA from a human pituitary gland-derived
CC cDNA library. The rat PACAP receptor gene was cloned by constructing
CC primers based on the homology between the PACAP receptor and the
CC vasoactive intestinal peptide (VIP) proteins. The probe was used to
CC obtain clones from a rat brain cDNA library. The PACAP receptor or
CC fragments thereof may be used for the diagnosis of neuropathy such as
CC Alzheimer's disease. The DNA encoding the PACAP receptor can be used for
CC gene therapy. Compounds isolated using the screening assay can be used
CC for treating neuropathy.

XX XX

SQ Sequence 76 AA;

Query Match 74.1%; Score 60; DB 15; Length 76;
Best Local Similarity 68.8%; Pred. No. 0.0031;
Matches 11; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

Qy 1 YC--NGEVAEVRKKMW 14

```

RESULT 11
R92277 standard; Protein; 591 AA.
XX
AC R92277;
XX
DT 18-MAY-1996 (first entry)
XX
DE Rat bone PTH/PTHrP receptor.
XX
KW Parathyroid hormone; receptor; parathormone; PTH;
KW parathyroid hormone-related protein; PTHrP; calcium; homeostasis;
KW hypercalcaemia; hypocalcaemia; cancer.
XX
OS Rattus sp.
XX
FH Key Location/Qualifiers
FT Region 1..192
FT /label= Extracellular_region
FT Region 193..211
FT /label= Transmembrane_region
FT Region 212..221
FT /label= Intracellular_region
FT Region 222..240
FT /label= Transmembrane_region
FT Region 241..299
FT /label= Extracellular_region
FT Region 300..316
FT /label= Transmembrane_region
FT Region 317..325
FT /label= Intracellular_region
FT Region 326..342
FT /label= Transmembrane_region
FT Region 343..364
FT /label= Extracellular_region
FT Region 365..383
FT /label= Transmembrane_region
FT Region 384..408
FT /label= Intracellular_region
FT Region 409..428
FT /label= Transmembrane_region
FT Region 429..444
FT /label= Intracellular_region
FT Region 445..463
FT /label= Transmembrane_region
FT Region 464..591
FT /label= Intracellular_region
XX
FN US5494806-A.
XX
PD 27-FEB-1996.
XX
PF 05-APR-1991; 91US-0681702.
XX
PR 06-APR-1992; 92US-0864475.
PR 05-APR-1991; 91US-0681702.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
PI Schipani E, Segre GV;
XX
WPI; 1996-139028/14.
DR N-PSDB; T15947.
XX
DT DNA encoding vertebrate parathyroid hormone receptor - useful for
FT diagnosis and treatment of e.g. hypercalcaemia, hypo-calcaemia,
PT cancer etc.
PT
PS Claim 1; Fig 3A-3E; 64pp; English.
XX
CC A rat parathyroid hormone/parathyroid hormone-related protein

```

```

CC (PTH/PTHrP) receptor (R92277) is encoded by cDNA clone R15B
CC (T15947) isolated from a rat osteosarcoma ROS 17/2.8 cDNA library.
CC The receptor is a G-protein linked receptor having 7 transmembrane
CC domains. It induces an increase in intracellular cAMP and calcium
CC upon challenge with PTH or PTHrP. Recombinant receptor can be
CC produced in vector/host cell systems and used in the treatment,
CC diagnosis or prognosis of e.g. hypercalcaemia and hypocalcaemia,
CC to screen for (ant)agonists and to raise antibodies. Host cells
CC expressing the receptor are used for diagnostic measurement of PTH
CC serum levels.
XX
SQ Sequence 591 AA;
XX
Query Match 82.7%; Score 67; DB 17; Length 591;
Best Local Similarity 71.4%; Pred. No. 0.0016;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Oy 1 YCNGEVQAEVKMW 14
Db 461 fcngvqaeirksw 474
:|||||::|
fcngvqaeirksw 474
RESULT 12
W73316
ID W73316 standard; Protein; 591 AA.
XX
AC W73316;
XX
DT 08-FEB-1999 (first entry)
XX
DE Parathyroid hormone receptor R15B.
XX
KW Parathyroid hormone receptor; PTH receptor; antibody; therapy;
KW PTH-related hypercalcaemia; rat.
XX
OS Rattus sp.
XX
FN US5840853-A.
XX
PD 24-NOV-1998.
XX
PF 06-JUN-1995; 95US-0471494.
XX
PR 06-APR-1992; 92US-0864475.
PR 05-APR-1991; 91US-0681702.
PR 06-JUN-1995; 95US-0471494.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
PI Schipani E, Segre GV;
XX
WPI; 1999-034124/03.
DR N-PSDB; V08390.
XX
PT Antibody to parathyroid hormone receptor - for diagnostic or
PT therapeutic use
XX
PS Claim 6; Fig 3; 63pp; English.
XX
CC This sequence represents the rat parathyroid hormone (PTH) receptor
CC R15B, which is targeted by the antibody of the invention. The antibody
CC of the invention is immunoreactive with naturally occurring human, rat or
CC opossum PTH receptor. The antibody is useful for treating disorders
CC characterised by overstimulation of PTH receptors by their ligand and for
CC the diagnosis of PTH-related hypercalcaemia.
XX
SQ Sequence 591 AA;
XX
Query Match 82.7%; Score 67; DB 20; Length 591;
Best Local Similarity 71.4%; Pred. No. 0.0016;

```

XX A human parathyroid hormone/parathyroid hormone-related protein
 CC (PTH/PTHrP) receptor (R92278) is encoded by cDNA clone HK-1
 CC (T15948) isolated from a human kidney cDNA library. The receptor
 CC induces an increase in intracellular cAMP and intracellular free
 CC calcium when challenged by PTH or PTHrP. Recombinant receptor can
 CC be produced in vector/host cell systems and used in the treatment,
 CC diagnosis or prognosis of e.g. hypercalcaemia and hypocalcaemia, to
 CC screen for (ant)agonists and to raise antibodies. Host cells
 CC expressing the receptor can be used for diagnostic measurement of
 CC PTH serum levels.

XX
 XX
 SQ Sequence 593 AA;

Query Match 86.4%; Score 70; DB 17; Length 593;
 Best Local Similarity 78.6%; Pred. No. 0.0005;
 Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YCNGEVQAEVKKMW 14
 :|||||:| |
 Db 461 fcngvqaeikksw 474

RESULT 9
 W73317
 ID W73317 standard; Protein; 593 AA.
 AC W73317;
 XX
 XX
 DT 08-FEB-1999 (first entry)
 XX
 DE Human Parathyroid hormone receptor.
 XX
 KW Parathyroid hormone receptor; PTH receptor; antibody; therapy;
 KW PTH-related hypercalcaemia; human.
 XX
 OS Homo sapiens.
 XX
 PN US5840853-A.
 XX
 PD 24-NOV-1998.
 XX
 PF 06-JUN-1995; 95US-0471494.
 XX
 PR 06-APR-1992; 92US-0864475.
 PR 05-APR-1991; 91US-0681702.
 PR 06-JUN-1995; 95US-0471494.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;
 XX
 DR WPI; 1999-034124/03.
 DR N-PSDB; V08391.
 XX
 PT Antibody to parathyroid hormone receptor - for diagnostic or
 PT therapeutic use
 XX
 PS Claim 7; Fig 6; 63pp; English.
 XX
 CC This sequence represents the human parathyroid hormone (PTH) receptor
 CC which is targeted by the antibody of the invention. The antibody
 CC of the invention is immunoreactive with naturally occurring human, rat or
 CC opossum PTH receptor. The antibody is useful for treating disorders
 CC characterised by overstimulation of PTH receptors by their ligand and for
 CC the diagnosis of PTH-related hypercalcaemia.

XX
 XX
 SQ Sequence 593 AA;

Query Match 86.4%; Score 70; DB 20; Length 593;

Best Local Similarity 78.6%; Pred. No. 0.0005;
 Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YCNGEVQAEVKKMW 14
 :|||||:| |
 Db 461 fcngvqaeikksw 474

RESULT 10
 R27706
 ID R27706 standard; Protein; 591 AA.
 XX
 AC R27706;
 XX
 DT 16-MAR-1993 (first entry)
 XX
 DE Rat bone PTH/PTHrP receptor clone R15B prod.
 XX
 KW Parathyroid hormone; related protein; calcium; antagonist;
 KW antibodies; hypercalcaemia.
 XX
 OS Rattus rattus.
 XX
 PN W09217602-A.
 XX
 PD 15-OCT-1992.
 XX
 PF 06-APR-1992; 92WO-US02821.
 XX
 PR 05-APR-1991; 91US-0681702.
 PR 06-APR-1992; 92US-0864475.
 XX
 PA (GEHO) GEN HOSPITAL CORP OFFICE TECHNOLOGY.
 XX
 PI Abou-samra A, Juppner H, Kronenberg HM, Potts JT, Schipani E;
 PI Segre GV;
 XX
 DR WPI; 1992-366271/44.
 DR N-PSDB; Q29606.
 XX
 PT New DNA encoding parathyroid hormone receptor, DNA and antibodies
 PT - for (differential) diagnosis of hypercalcaemia, and diagnosis
 PT and treatment of tumours
 XX
 PS Claim 20; Fig 3; 91pp; English.
 XX
 CC The rat bone parathyroid hormone/parathyroid hormone related
 CC protein (PTH/PTHrP) receptor protein sequence was deduced from
 CC clone R15B obtd. by screening a rat osteosarcoma (ROS) cell cDNA library
 CC to isolate those expressing functionally intact PTH/PTHrP receptor
 CC proteins, performed according to Gearing et al., (EMBO J. 8: 3676,
 CC 1989), by identifying colonies capable of binding a suitable radio-
 CC labelled ligand. The protein may be used in a therapeutic compsn. to
 CC inhibit activation of PTH or PTHrP and thus reduce the level of calcium
 CC in the blood. Cpds. capable of competing with PTH or PTHrP for binding
 CC can be identified using the protein and DNAs homologous to PTH DNA can
 CC be identified using fragments of the clone as probes. The sequence
 CC may be used for the prodn. of antibodies useful for the treatment,
 CC classification, prognosis and/or treatment of disorders related to
 CC the interaction between a cell receptor and a ligand such as in
 CC hypercalcaemia. See also R27704-16.

XX
 SQ Sequence 591 AA;

Query Match 82.7%; Score 67; DB 13; Length 591;
 Best Local Similarity 71.4%; Pred. No. 0.0016;
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YCNGEVQAEVKKMW 14
 :|||||:| |
 Db 461 fcngvqaeikksw 474

KW hypercalcaemia; hypocalcaemia; cancer; opossum.
 OS Didelphis virginiana.
 PN US5494806-A.
 XX 27-FEB-1996.
 XX
 PF 05-APR-1991; 91US-0681702.
 XX
 PR 06-APR-1992; 92US-0864475.
 XX
 PR 05-APR-1991; 91US-0681702.
 XX
 PA (GEO) GEN HOSPITAL CORP.
 XX
 PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;
 XX
 DR WPI; 1996-139028/14.
 DR N-PSDB; T15946.
 XX
 PT DNA encoding vertebrate parathyroid hormone receptor - useful for
 PT diagnosis and treatment of e.g. hypercalcaemia, hypo-calcaemia,
 PT cancer etc.
 XX
 PS Claim 1; Fig 2A-2E; 64pp; English.
 XX
 CC Parathyroid hormone/parathyroid hormone-related protein (PTH/PTHrP)
 CC receptors (R92275 and R92776) are encoded by cDNA clones OK-H (T15945)
 CC and OK-O (T15946), respectively, isolated from an opossum kidney (OK)
 CC cell cDNA library. OK-H and OK-O may be the products of 2 separate
 CC genes or of a laboratory artifact. The receptor induces an increase
 CC in intracellular cAMP and calcium when challenged with PTH or PTHrP.
 CC Recombinant receptors can be produced in vector/host cell systems and
 CC used in the treatment, diagnosis or prognosis of e.g. hypercalcaemia
 CC and hypocalcaemia, to screen for (ant)agonists and to raise antibodies.
 CC Host cells expressing the receptor can be used for diagnostic
 CC measurement of PTH serum levels.
 XX
 SQ Sequence 585 AA;
 XX
 Query Match 86.4%; Score 70; DB 17; Length 585;
 Best Local Similarity 78.6%; Pred. No. 0.00051.
 Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 YCNGEVQAEVKKMW 14
 :|||||||:|
 Db 455 fnggevqaeikksw 468
 XX
 RESULT 7
 W73315
 ID W73315 standard; Protein; 585 AA.
 XX
 AC W73315;
 XX
 DT 08-FEB-1999 (first entry)
 XX
 DE Parathyroid hormone receptor OK-O.
 XX
 KW Parathyroid hormone receptor; PTH receptor; antibody; therapy;
 KW PTH-related hypercalcaemia; opossum.
 XX
 OS Didelphis virginiana.
 XX
 PN US5840853-A.
 XX
 PD 24-NOV-1998.
 XX
 PF 06-JUN-1995; 95US-0471494.
 XX
 PR 06-APR-1992; 92US-0864475.

PR 05-APR-1991; 91US-0681702.
 PR 06-JUN-1995; 95US-0471494.
 XX
 PA (GEO) GEN HOSPITAL CORP.
 XX
 PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;
 XX
 DR WPI; 1999-034124/03.
 DR N-PSDB; V08389.
 XX
 PT Antibody to parathyroid hormone receptor - for diagnostic or
 PT therapeutic use
 XX
 PS Claim 6; Fig 2; 63pp; English.
 XX
 CC This sequence represents the opossum parathyroid hormone (PTH) receptor
 CC OK-O, which is targeted by the antibody of the invention. The antibody
 CC of the invention is immunoreactive with naturally occurring human, rat or
 CC opossum PTH receptor. The antibody is useful for treating disorders
 CC characterised by overstimulation of PTH receptors by their ligand and for
 CC the diagnosis of PTH-related hypercalcaemia.
 XX
 SQ Sequence 585 AA;
 XX
 Query Match 86.4%; Score 70; DB 20; Length 585;
 Best Local Similarity 78.6%; Pred. No. 0.0005;
 Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 YCNGEVQAEVKKMW 14
 :|||||||:|
 Db 455 fnggevqaeikksw 468
 XX
 RESULT 8
 R92278
 ID R92278 standard; Protein; 593 AA.
 XX
 AC R92278;
 XX
 DT 18-MAY-1996 (first entry)
 XX
 DE Human kidney PTH/PTHrP receptor.
 XX
 KW Parathyroid hormone; receptor; parathormone; PTH;
 KW parathyroid hormone-related protein; PTHrP; calcium; homeostasis;
 KW hypercalcaemia; hypocalcaemia; cancer.
 XX
 OS Homo sapiens.
 XX
 PN US5494806-A.
 XX
 PD 27-FEB-1996.
 XX
 PF 05-APR-1991; 91US-0681702.
 XX
 PR 06-APR-1992; 92US-0864475.
 XX
 PR 05-APR-1991; 91US-0681702.
 XX
 PA (GEO) GEN HOSPITAL CORP.
 XX
 PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;
 XX
 DR WPI; 1996-139028/14.
 DR N-PSDB; T15948.
 XX
 PT DNA encoding vertebrate parathyroid hormone receptor - useful for
 PT diagnosis and treatment of e.g. hypercalcaemia, hypo-calcaemia,
 PT cancer etc.
 XX
 PS Claim 1; Fig 6A-6G; 64pp; English.

Query Match 86.4%; Score 70; DB 17; Length 515;
 Best Local Similarity 78.6%; Pred. No. 0.00044;
 Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 YCNGEVQAEVKKMW 14
 Db 455 fcngvqaeikksw 468
 :|||||:| |

RESULT 4
 W73314
 ID W73314 standard; Protein; 515 AA.
 XX AC W73314;
 XX DF 08-FEB-1999 (first entry)
 XX DE Parathyroid hormone receptor OK-H.
 XX KW Parathyroid hormone receptor; PTH receptor; antibody; therapy;
 KW PTH-related hypercalcaemia; opossum.
 XX OS Didelphis virginiana.
 XX PN US5840853-A.
 XX PD 24-NOV-1998.
 XX PF 06-JUN-1995; 95US-0471494.
 XX PR 06-APR-1992; 92US-0864475.
 PR 05-APR-1991; 91US-0681702.
 XX PR 06-JUN-1995; 95US-0471494.
 XX PA (GEHO) GEN HOSPITAL CORP.
 XX PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;
 DR WPI; 1999-034124/03.
 DR N-PSDB; V08388.
 XX Antibody to parathyroid hormone receptor - for diagnostic or
 PT therapeutic use
 PS Claim 6; Fig 1; 63pp; English.
 CC This sequence represents the opossum parathyroid hormone (PTH) receptor
 CC OK-H, which is targeted by the antibody of the invention. The antibody
 CC of the invention is immunoreactive with naturally occurring human, rat or
 CC opossum PTH receptor. The antibody is useful for treating disorders
 CC characterised by overstimulation of PTH receptors by their ligand and for
 CC the diagnosis of PTH-related hypercalcaemia.
 XX SQ Sequence 515 AA;

Query Match 86.4%; Score 70; DB 20; Length 515;
 Best Local Similarity 78.6%; Pred. No. 0.00044;
 Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 YCNGEVQAEVKKMW 14
 Db 455 fcngvqaeikksw 468
 :|||||:| |

RESULT 5
 R27705
 ID R27705 standard; Protein; 585 AA.
 XX AC R27705;
 XX DT 16-MAR-1993 (first entry)

Query Match 86.4%; Score 70; DB 13; Length 585;
 Best Local Similarity 78.6%; Pred. No. 0.0005;
 Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 YCNGEVQAEVKKMW 14
 Db 455 fcngvqaeikksw 468
 :|||||:| |

RESULT 6
 R92276
 ID R92276 standard; Protein; 585 AA.
 XX AC R92276;
 XX DT 18-MAY-1996 (first entry)
 XX DE Opossum kidney PTH/PTHrP receptor.
 KW Parathyroid hormone; receptor; parathormone; PTH;
 KW parathyroid hormone-related protein; PTHrP; calcium; homeostasis;

XX DE Opossum kidney PTH/PTHrP receptor prod. from clone OK-O.
 XX KW Parathyroid hormone; related protein; calcium; antagonist;
 KW antibodies; hypercalcaemia.
 XX OS Didelphis virginiana.
 XX PN W09217602-A.
 XX PD 15-OCT-1992.
 XX PF 06-APR-1992; 92WO-US02821.
 XX PR 05-APR-1991; 91US-0681702.
 PR 06-APR-1992; 92US-0864475.
 XX PA (GEHO) GEN HOSPITAL CORP OFFICE TECHNOLOGY.
 XX PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT, Schipani E;
 PI Segre GV;
 XX DR WPI; 1992-366271/44.
 DR N-PSDB; Q29605.
 XX New DNA encoding parathyroid hormone receptor, DNA and antibodies
 PT - for (differential) diagnosis of hypercalcaemia, and diagnosis
 PT and treatment of tumours
 XX FS Disclosure; Fig 2; 91pp; English.
 XX CC The parathyroid hormone/parathyroid hormone related protein (PTH/PTHrP)
 CC receptor protein sequence was deduced from the DNA sequence of the
 CC clone OK-O, isolated from opossum kidney (OK) cells. The clone OK-O
 CC is identical to the OK-H clone except at the C-terminal tail as OK-O
 CC encodes a 585 amino acid protein, OK-H encodes one of 515 amino acids.
 CC The difference is attributed to a single nucleotide deleted in the OK-H
 CC sequence causing a frame shift and an earlier stop codon. It is not
 CC known whether OK-O and OK-H represent prods. of two separate genes or
 CC are a laboratory artifact. The protein may be used in a therapeutic
 CC compsn. to inhibit activation of PTH or PTHrP and thus reduce the
 CC level of calcium in the blood. Cpd. capable of competing with PTH
 CC or PTHrP for binding can be identified using the protein prod. and
 CC DNAs homologous to PTH DNA can be identified using fragments of the
 CC clone as probes. The sequence may be used for the prodn. of antibodies
 CC useful for the treatment, classification, prognosis and/or treatment of
 CC disorders related to the interaction between a cell receptor and a
 CC ligand such as in hypercalcaemia. See also R27704-16.
 XX SQ Sequence 585 AA;

Query Match 86.4%; Score 70; DB 13; Length 585;
 Best Local Similarity 78.6%; Pred. No. 0.0005;
 Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 YCNGEVQAEVKKMW 14
 Db 455 fcngvqaeikksw 468
 :|||||:| |

RESULT 6
 R92276
 ID R92276 standard; Protein; 585 AA.
 XX AC R92276;
 XX DT 18-MAY-1996 (first entry)
 XX DE Opossum kidney PTH/PTHrP receptor.
 KW Parathyroid hormone; receptor; parathormone; PTH;
 KW parathyroid hormone-related protein; PTHrP; calcium; homeostasis;

PT hyper-calcaemia, hypo- or hyper-phosphatemia, kidney stones, etc
 PS Claim 9; Fig 1A-E; 62pp; English.

CC A novel 7-transmembrane receptor (W12695) has been identified as a
 CC human G-protein parathyroid hormone (PTH) receptor, designated
 CC H12DG74. It shows 48.2% homology to the human PTH receptor. Its
 CC amino acid sequence was deduced from a cDNA clone (T59619) isolated
 CC from a human T cell lymphoma tissue cDNA library. Recombinant
 CC H12DG74 can be produced in transformed host cells and used to
 CC screen for (ant)agonist cpds. Agonists may be used to prevent or
 CC treat e.g. hypocalcaemia, hyperphosphataemia, hypoparathyroidism
 CC and chronic tetany by stimulating an increase in serum calcium
 CC levels. Antagonists can be used to inhibit the receptor e.g. for
 CC the treatment of osteoporosis, hypercalcaemia, hypoparathyroidism,
 CC hypophosphataemia, kidney stone, nephrolithiasis.

XX Sequence 541 AA;

Query Match 100.0%; Score 81; DB 18; Length 541;
 Best Local Similarity 100.0%; Pred. NO. 6.3e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YCNGEVQAEVKKMW 14
 |||||

DB 415 ycngeqvqaeikkmw 428

RESULT 2

R27704
 ID R27704 standard; Protein; 515 AA.

XX R27704;

XX 16-MAR-1993 (first entry)

DE Opossum kidney PTH/PTHrP receptor prepd. from clone OK-H.

XX Parathyroid hormone; related protein; calcium; antagonist;
 KW antibodies; hypercalcaemia.

XX Didelphis virginiana.

XX WO9217602-A.

XX 15-OCT-1992.

XX 06-APR-1992; 92WO-US02821.

XX 05-APR-1991; 91US-0681702.

XX 06-APR-1992; 92US-0864475.

XX (GEHO) GEN HOSPITAL CORP OFFICE TECHNOLOGY.

XX Abou-samra A, Juppner H, Kronenberg HM, Potts JT, Schipani E;
 PI Segre GV;

XX WPI; 1992-366271/44.

XX N-PSDB; Q29604.

XX New DNA encoding parathyroid hormone receptor, DNA and antibodies
 PT - for (differential) diagnosis of hypercalcaemia, and diagnosis
 PT and treatment of tumours

XX Disclosure; Fig 1; 91pp; English.

XX The parathyroid hormone/parathyroid hormone related protein (PTH/PTHrP)
 CC receptor protein sequence was deduced from the DNA sequence of the
 CC clone OK-H, isolated from opossum kidney (OK) cells. The protein
 CC may be used in a therapeutic compsn. to inhibit activation of PTH or
 CC PTHrP and thus reduce the level of calcium in the blood. Cpds.
 CC capable of competing with PTH or PTHrP for binding can be identified

CC using the protein prod. and DNAs homologous to PTH DNA can be
 CC identified using fragments of the clone as probes. The protein
 CC may be used for the prodn. of antibodies useful for the treatment,
 CC classification, prognosis and/or treatment of disorders related to
 CC the interaction between a cell receptor and a ligand such as in
 CC hypercalcaemia. See also R27705-16.

XX Sequence 515 AA;

Query Match 86.4%; Score 70; DB 13; Length 515;
 Best Local Similarity 78.6%; Pred. NO. 0.00044;
 Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 YCNGEVQAEVKKMW 14
 :|||||:|

DB 455 fcngvqaeikkmw 468

RESULT 3

R92275
 ID R92275 standard; Protein; 515 AA.

XX R92275;

XX 18-MAY-1996 (first entry)

DE Opossum kidney PTH/PTHrP receptor.

XX Parathyroid hormone; receptor; parathormone; PTH;
 KW parathyroid hormone-related protein; PTHrP; calcium; homeostasis;
 KW hypercalcaemia; hypocalcaemia; cancer; opossum.

XX Didelphis virginiana.

XX US5494806-A.

XX 27-FEB-1996.

XX 05-APR-1991; 91US-0681702.

XX 06-APR-1992; 92US-0864475.

XX 05-APR-1991; 91US-0681702.

XX (GEHO) GEN HOSPITAL CORP.

XX Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;

XX WPI; 1996-139028/14.

XX N-PSDB; T15945.

XX DNA encoding vertebrate parathyroid hormone receptor - useful for
 PT diagnosis and treatment of e.g. hypercalcaemia, hypo-calcaemia,
 PT cancer etc.

XX Claim 1; Fig 1A-1E; 64pp; English.

XX Parathyroid hormone/parathyroid hormone-related protein (PTH/PTHrP)
 CC receptors (R92275 and R92776) are encoded by cDNA clones OK-H (T15945)
 CC and OK-O (T15946), respectively, isolated from an opossum kidney (OK)
 CC cell cDNA library. OK-H and OK-O may be the products of 2 separate
 CC genes or of a laboratory artifact. The receptor induces an increase
 CC in intracellular cAMP and calcium when challenged with PTH or PTHrP.
 CC Recombinant receptors can be produced in vector/host cell systems and
 CC used in the treatment, diagnosis or prognosis of e.g. hypercalcaemia
 CC and hypocalcaemia, to screen for (ant)agonists and to raise antibodies.
 CC Host cells expressing the receptor can be used for diagnostic
 CC measurement of PTH serum levels.

XX Sequence 515 AA;

OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; *Drosophila*.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY.
 RX MEDLINE; 20196006.
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003639; AAF53247.1; -;
 DR FLYBASE; FBgn0032474; CG9828.
 DR INTERPRO; IPR001305; -;
 DR INTERPRO; IPR001623; -;
 DR INTERPRO; IPR002939; -;
 DR INTERPRO; IPR003095; -;
 DR PFAM; PF00226; DnaJ; 1.
 DR PFAM; PF00684; DnaJ_CXXCXXG; 1.
 DR PFAM; PF01556; DnaJ_C; 1.
 DR PRINTS; PR00625; DNAJPROTEIN.
 DR PROSITE; PS00636; DNAJ_1; 1.
 DR PROSITE; PS50076; DNAJ_2; 1.
 SQ SEQUENCE 389 AA; 43938 MW; FC847DC8305FAE19 CRC64;

Query Match 52.6%; Score 41; DB 5; Length 389;
 Best Local Similarity 53.3%; Pred. No. 28;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AVGHDTKQYKRLAK 15
 | : |||||
 Db ..16 ATDEEKKNYKRLAK 30

Search completed: November 8, 2000, 08:56:12
 Job time: 523 sec

```
ID O54981 PRELIMINARY; PRT; 543 AA.
AC O54981;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE HSP70/HSP90 ORGANIZING PROTEIN.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Ovary;
RA Heine H., Delude R.L., Monks B., Golenbock D.T.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF039202; AAB94760.1; -.
DR INTERPRO: IPR001440; -.
DR PFAM: PF00515; TPR; 4.
SQ SEQUENCE 543 AA; 62651 MW; 4B6A9C17EEFBF287 CRC64;

Query Match 53.8%; Score 42; DB 11; Length 543;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VGHDTKQYRKLA 15
|| : | : || : ||
Db 288 VGRENEDYRQIAK 301

RESULT 12
Q60864 PRELIMINARY; PRT; 543 AA.
AC Q60864;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE EXTENDIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-M27 LEWIS LUNG CARCINOMA;
RA Blatch G.L., Lasse M., Takatori T., Gandhi T., Kundra V.,
Zetter B.R.;
RL Proc. Am. Assoc. Cancer Res. 36:68-68(1995).
DR EMBL: U27830; AAC53267.1; -.
DR INTERPRO: IPR001440; -.
DR PFAM: PF00515; TPR; 4.
SQ SEQUENCE 543 AA; 62582 MW; B737FBA92B198D6C CRC64;

Query Match 53.8%; Score 42; DB 11; Length 543;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VGHDTKQYRKLA 15
|| : | : || : ||
Db 288 VGRENEDYRQIAK 301

RESULT 13
O35814 PRELIMINARY; PRT; 543 AA.
AC O35814;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE P60 PROTEIN.
GN HOP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
RX MEDLINE; 98187623.
RA Demand J., Luters J., Hoehfeld J.;
RT "The carboxy-terminal domain of Hsc70 provides binding sites for a
distinct set of chaperone cofactors.";
RL Mol. Cell. Biol. 18:2023-2028(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
RA Hoehfeld J.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y15068; CAA75351.1; -.
DR INTERPRO: IPR001440; -.
DR PFAM: PF00515; TPR; 4.
SQ SEQUENCE 543 AA; 62570 MW; D8313F43BFB7EB73 CRC64;

Query Match 53.8%; Score 42; DB 11; Length 543;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VGHDTKQYRKLA 15
|| : | : || : ||
Db 288 VGRENEDYRQIAK 301

RESULT 14
Q9SYD4 PRELIMINARY; PRT; 375 AA.
AC Q9SYD4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE FLIM15, 18 PROTEIN.
GN FLIM15, 18.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; Rosidae; eurosids II; Brassicales;
OC Brassicaceae; Arabidopsids.
RN [1]
RP SEQUENCE FROM N.A.
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
Alcafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
Luros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC006085; AAC30643.1; -.
DR INTERPRO: IPR001810; -.
SQ SEQUENCE 375 AA; 44207 MW; 809FCB34CE0AE105 CRC64;

Query Match 52.6%; Score 41; DB 10; Length 375;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VGHDTKQYRKLA 13
|| : | : || : |
Db 138 IGYDNKKQYKML 149

RESULT 15
Q9VK35 PRELIMINARY; PRT; 389 AA.
AC Q9VK35;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CG9828 PROTEIN.
GN CG9828.
```

```

RX MEDLINE; 94150718.
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstein L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR EMBL; Z75541; CAA99855.1; -.
DR INTERPRO; IPR001374; -.
DR INTERPRO; IPR001410; -.
DR INTERPRO; IPR001650; -.
DR PFAM; PF00271; helicase_C; 1.
DR PFAM; PF01424; R3H; 1.
SQ SEQUENCE 1425 AA; 162466 MW; 2A04352CC1931199E CRC64;

Query Match 55.1%; Score 43; DB 5; Length 1425;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 4 HDTKQYRKLA 14
Db 35 NDTKREYRVA 45
:||||:|::|

RESULT 10
Q17438 PRELIMINARY; PRT; 401 AA.
AC Q17438;
DT 01-JAN-1999 (TrEMBLrel. 09, Created)
DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE HYPOTHETICAL 46.6 KDA PROTEIN B0035.14 IN CHROMOSOME IV.
GN B0035.14;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RL White S.;
RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS A DNAJ-LIKE DOMAIN.
DR EMBL; Z73102; CAA97416.1; -.
DR HSSP; P25685; 1HDJ.
DR WORMPEP; B0035.14; CE05169.
DR INTERPRO; IPR001623; -.
DR PFAM; PF00226; DnaJ; 1.
DR PROSITE; PS00636; DNAJ_1; 1.
DR PROSITE; PS50076; DNAJ_2; 1.
KW Hypothetical protein; Chaperone.
FT DOMAIN 134 203 DNAJ-LIKE.
SQ SEQUENCE 401 AA; 46582 MW; 74E2170B3BADBD5A CRC64;

Query Match 53.8%; Score 42; DB 5; Length 401;
Best Local Similarity 80.08%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 DTRKQYRKLA 14
Db 152 DIRKQYRKLA 161
| ||| |||

RESULT 11
Q054981

```

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Beson P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler G., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali B., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel M.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of *Drosophila melanogaster*."
RA Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS CSP2 AND CSP3).
RC STRAIN-BERLIN;
RX MEDLINE; 99015937.
RA Eberle K.K., Zinsmaier K.E., Buchner S., Gruhn M., Jenni M.,
RA Arnold C., Leibold C., Reisch D., Walter N., Hafen E., Hofbauer A.,
RA Pflugfelder G.O., Buchner E.;
RA "Wide distribution of the cysteine string proteins in *Drosophila*
RT tissues revealed by targeted mutagenesis."
RL Cell Tissue Res. 294:203-217(1998).
CC -1- ALTERNATIVE PRODUCTS: THREE ISOFORMS; ISOFORM CSP1 (OR CSP32)
CC (SHOWN HERE), ISOFORM CSP2 AND ISOFORM CSP3 (OR CSP29); ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC EMBL; AE0033597; AAF51816.1; -.
DR EMBL; AE0033597; AAF51817.1; -.
DR EMBL; AF057167; AAD09430.1; -.
DR EMBL; AF057167; AAD09431.1; -.
DR HSP; P25685; 1HDJ.
DR FLYBASE; FBgn0004179; Csp.
DR INTERPRO; IPR001450; -.
DR INTERPRO; IPR001623; -.
DR INTERPRO; IPR003095; -.
DR PFAM; PF00226; DnaJ; 1.
DR PRINTS; PR00625; DnaJPROTEIN.
DR PROSITE; PS00198; 4FEAS_FERREDOXIN; 1.
DR PROSITE; PS00636; DnaJ_1; 1.
DR PROSITE; PS0076; DnaJ_2; 1.
KW Alternative splicing.
FT VARSPPLIC 154 174 MISSING (IN ISOFORM CSP3 AND ISOFORM
FT CSP2).
FT VARSPPLIC 243 249 DMVNOKY -> GI (IN ISOFORM CSP3).
SQ SEQUENCE 249 AA; 26897 MW; 6EF37B080985F3903 CRC64;

Query Match 57.7%; Score 45; DB 5; Length 249;
Best Local Similarity 64.3%; Pred. No. 3.6;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AVGHTRKQYRKLA 14
| | | | |
Db 28 ATGDDIKTKYRKLA 41

RESULT 6
Q9UQY8 PRELIMINARY; PRT; 56 AA.
ID Q9UQY8
AC Q9UQY8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE PUTATIVE MAT-2 MATING TYPE PROTEIN (FRAGMENT)
OS Colletotrichum graminicola (Anthracoze fungus)
OS (Glomerella graminicola).
OC Eukaryota; Fungi; Ascomycota; Phyllachorales; Phyllachoraceae;
OC Glomerella.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-M5.001, AND M1.001;
RA Vallancourt L.J., Du M., Wang J., Rollins J., Hanau R.;
RT "Genetic analysis of cross-fertility between two self-sterile strains
RT of *Glomerella graminicola*."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF204961; AAF23339.1; -.
DR EMBL; AF204960; AAF23338.1; -.
DR INTERPRO; IPR000910; -.
DR PFAM; PF00505; HMG_box; 1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6688 MW; B2BBA04C95819EAE CRC64;

Query Match 55.1%; Score 43; DB 3; Length 56;
Best Local Similarity 53.3%; Pred. No. 1.7;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AVGHTRKQYRKLA 15
| | | | |
Db 33 AESHEVREKYLAK 47

RESULT 7
O57671 PRELIMINARY; PRT; 126 AA.
ID O57671
AC O57671;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE PARATHYROID HORMONE RECEPTOR (FRAGMENT).
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagridae; Meleagris.
RN [1]
RP SEQUENCE FROM N.A.
RA Hsu C., You S., el Halawani M.E., Foster D.N.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U94326; AAB93893.1; -.
DR INTERPRO; IPR000832; -.
DR PFAM; PF00002; 7tm_2; 1.
DR NON_TER 1
DR NON_TER 126
FT NON_TER 126
SQ SEQUENCE 126 AA; 14515 MW; 0F381BD8094A1A77 CRC64;

Query Match 55.1%; Score 43; DB 13; Length 126;
Best Local Similarity 72.7%; Pred. No. 3.9;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRKQYRKLA 15
| | | | |
Db 104 DSRQYRKLLK 114

RESULT 8
Q19546

```

ID Q9PWB7 PRELIMINARY; PRT; 575 AA.
AC Q9PWB7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE PARATHYROID HORMONE TYPE-2 RECEPTOR PRECURSOR.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprinidae; Rasbora; Danio.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-KIDNEY;
RX MEDLINE; 99367425.
RA Rubin D.A., Hellman P., Zon L.I., Lobb C.J., Bergwitz C., Juppner H.;
RT "A G protein-coupled receptor from zebrafish is activated by human
RT parathyroid hormone and not by human or teleost parathyroid hormone-
RT related peptide. Implications for the evolutionary conservation of
RT calcium-regulating peptide hormones."
RL J. Biol. Chem. 274:23035-23042(1999).
DR EMBL; AF132082; AAD51908.1; -.
DR INTERPRO; IPR000832; -.
DR INTERPRO; IPR001879; -.
DR PFAM; PF00002; 7tm.2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 575 AA; 64244 MW; 888F1C4DDB3A14DC CRC64;

Query Match 74.4%; Score 58; DB 13; Length 575;
Best Local Similarity 91.7%; Pred. No. 0.049;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 HDTKQYRKLLK 15
Db 379 YDTRQYRKLLK 390

ID Q9PVD3 PRELIMINARY; PRT; 536 AA.
AC Q9PVD3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE PARATHYROID HORMONE RECEPTOR PTHL.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprinidae; Rasbora; Danio.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-KIDNEY;
RX MEDLINE; 99367425.
RA Rubin D.A., Juppner H.;
RT "Zebrafish express the Common Parathyroid Hormone/Parathyroid Hormone-
RT related Peptide Receptor (PTHr) and a Novel Receptor (PTH3R) That Is
RT Preferentially Activated by Mammalian and Fuguish Parathyroid
RT Hormone-related Peptide."
RL J. Biol. Chem. 274:28185-28190(1999).
DR EMBL; AF132084; AAF01265.1; -.
DR INTERPRO; IPR000832; -.
DR INTERPRO; IPR001879; -.
DR INTERPRO; IPR002170; -.
DR PFAM; PF00002; 7tm.2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PRINTS; PR00393; PTHHORMONER.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 536 AA; 61454 MW; 653B0A56DB2C8FA9 CRC64;

```

```

Query Match 60.3%; Score 47; DB 13; Length 536;
Best Local Similarity 81.8%; Pred. No. 3.6;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 DTRKQYRKLLK 15
Db 352 DTRQYRKLLK 362

ID Q9TU31 PRELIMINARY; PRT; 595 AA.
AC Q9TU31;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE PARATHYROID HORMONE RECEPTOR-1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-KIDNEY;
RA Smock S.L., Vogt G.A., Castleberry T.A., Lu B., Owen T.A.;
RT "Molecular cloning and functional characterization of the canine
RT parathyroid hormone receptor-1 (PTH1)."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF167095; AAD55938.1; -.
DR INTERPRO; IPR000832; -.
DR INTERPRO; IPR001879; -.
DR INTERPRO; IPR002170; -.
DR PFAM; PF00002; 7tm.2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PRINTS; PR00393; PTHHORMONER.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 595 AA; 66308 MW; 09568FCF38D4D258 CRC64;

Query Match 60.3%; Score 47; DB 6; Length 595;
Best Local Similarity 81.8%; Pred. No. 4;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 DTRKQYRKLLK 15
Db 397 DTRQYRKLLK 407

ID O61664 PRELIMINARY; PRT; 249 AA.
AC O61664; O61665; Q9VNV1;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CYSTEINE-STRING PROTEIN.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS CSPI AND CSP3).
RC STRAIN-BERKELEY;
RX MEDLINE; 20196006.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:56:10 ; Search time 152.43 Seconds
(without alignments)
9.189 Million cell updates/sec

Title: US-09-236-468a-2_COPY_349_363
Perfect score: 78
Sequence: 1 AVGHDRKQYRKLA 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 297973

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72	92.3	169	11 Q9R1D4	Q9r1d4 mus musculus
2	58	74.4	575	13 Q9PWB7	Q9pwb7 brachydanio
3	47	60.3	536	13 Q9PVD3	Q9pvd3 brachydanio
4	47	60.3	595	6 Q9TU31	Q9tu31 canis fam1
5	45	57.7	249	5 Q61664	Q61664 drosophila
6	43	55.1	56	3 Q9UQV8	Q9uqy8 colletotric
7	43	55.1	126	13 Q57671	Q57671 meleagris g
8	43	55.1	986	5 Q19546	Q19546 caenorhabdi
9	43	55.1	1425	5 Q20644	Q20644 caenorhabdi
10	42	53.8	401	5 Q17438	Q17438 caenorhabdi
11	42	53.8	543	11 Q54981	Q54981 cricetus
12	42	53.8	543	11 Q60864	Q60864 mus musculus
13	42	53.8	543	11 Q35814	Q35814 rattus norv
14	41	52.6	375	10 Q9SYD4	Q9syd4 arabidopsis
15	41	52.6	389	5 Q9VK35	Q9vk35 drosophila
16	41	52.6	439	2 Q9Z4R6	Q9z4r6 eikenella c
17	41	52.6	542	13 Q9PVD2	Q9pvd2 brachydanio
18	40	51.3	56	4 Q99641	Q99641 homo sapien
19	40	51.3	167	6 Q29456	Q29456 bos taurus

20	40	51.3	198	6 Q29455	Q29455 bos taurus
21	40	51.3	204	8 Q47112	Q47112 kigelia sp.
22	40	51.3	331	2 Q9ZEJ5	Q9zej5 anabaena sp
23	40	51.3	333	2 Q9ZEK1	Q9zek1 anabaena va
24	40	51.3	719	2 Q69992	Q69992 streptomyce
25	40	51.3	1099	5 Q9VMS5	Q9vms5 drosophila
26	40	51.3	1487	5 Q9Y062	Q9y062 drosophila
27	40	51.3	1487	5 Q9VGI8	Q9vgi8 drosophila
28	40	51.3	2241	5 Q15850	Q15850 leishmania
29	39	50.0	320	5 Q16259	Q16259 caenorhabdi
30	39	50.0	387	1 Q9UXR9	Q9uxr9 methanosarc
31	39	50.0	567	10 Q9ZWF0	Q9zwf0 arabidopsis
32	39	50.0	603	13 Q9YI32	Q9yi32 carassius a
33	39	50.0	627	5 Q15811	Q15811 plasmodium
34	39	50.0	672	5 Q9U0J2	Q9u0j2 plasmodium
35	39	50.0	942	10 Q9ZUQ8	Q9zuq8 arabidopsis
36	39	50.0	1998	5 Q9V7A5	Q9v7a5 drosophila
37	38	48.7	150	5 Q9TZE4	Q9tze4 caenorhabdi
38	38	48.7	259	12 Q9YKK0	Q9ykk0 epizootic h
39	38	48.7	259	12 Q9WQ99	Q9wq99 silurus gla
40	38	48.7	259	12 Q9WPP7	Q9wpp7 frog virus
41	38	48.7	342	10 Q9SP09	Q9sp09 nicotiana t
42	38	48.7	412	4 Q60884	Q60884 homo sapien
43	38	48.7	412	11 Q35824	Q35824 rattus norv
44	38	48.7	412	11 Q9QVJ0	Q9qvj0 mus musculu
45	38	48.7	415	4 Q14711	Q14711 homo sapien

ALIGNMENTS

RESULT 1

Q9R1D4

ID Q9R1D4 PRELIMINARY; PRT: 169 AA.

AC Q9R1D4;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

DE PARATHYROID HORMONE TYPE-2 RECEPTOR PRECURSOR (FRAGMENT).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;

RA MEDLINE; 99367425.

RA Rubin D.A., Hellman P., Zon L.I., Lobb C.J., Bergwitz C., Juppner H.;

RT "A G protein-coupled receptor from zebrafish is activated by human parathyroid hormone and not by human or teleost parathyroid hormone-related peptide. Implications for the evolutionary conservation of calcium-regulating peptide hormones.";

RL J. Biol. Chem. 274:23035-23042(1999).

DR EMBL; AF132083; AAC51909.1; -.

DR INTERPRO; IPR000832; -.

DR PFAM; PF00002; 7tm.2; 1.

DR PRINTS; PR00249; GPCRSECRETIN.

KW Receptor.

FT NON_TER 169 169

FT NON_TER 169 169

SQ SEQUENCE 169 AA; 19674 MW; 748CC8231F1C69EA CRC64;

Query Match 92.3%; Score 72; DB 11; Length 169;
Best Local Similarity 93.3%; Pred. No. 5.3e-05;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AVGHDRKQYRKLA 15

Db 99 AVGHDRKQYRKLA 113

RESULT 2

Q9PWB7

RX MEDLINE; 95223109.
RA Mastrogiacom A., Gundersen C.B.;
RT "The nucleotide and deduced amino acid sequence of a rat cysteine
RL string protein.";
RN Brain Res. Mol. Brain Res. 28:12-18(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=RAT; TISSUE=BRAIN;
RX MEDLINE; 96188189.
RA Braun J.E., Scheller R.H.;
RT "Cysteine string protein, a DnaJ family member, is present on diverse
RL secretory vesicles";
RN Neuropharmacology 34:1361-1369(1995).
CC -!- FUNCTION: MAY HAVE AN IMPORTANT ROLE IN PRESYNAPTIC FUNCTION. MAY
CC BE INVOLVED IN CALCIUM-DEPENDENT NEUROTRANSMITTER RELEASE AT NERVE
CC ENDINGS.
CC -!- SUBUNIT: HOMODIMER (PROBABLE).
CC -!- TISSUE SPECIFICITY: BRAIN. PREDOMINANTLY ASSOCIATED WITH NERVE
CC ENDINGS AND SYNAPTIC VESICLES.
CC -!- PTM: FATTY ACYLATED.
CC -!- SIMILARITY: CONTAINS A DNAJ-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39320; AAA81372.1; --
DR EMBL; S81917; AAB36303.1; --
DR EMBL; AF032115; AAB87080.1; --
DR HSPP; P25685; IHDJ.
DR MGD; MGI:892995; CSP.
DR INTERPRO; IPR001623; --
DR PFAM; PF00226; DnaJ; 1.
DR PROSITE; PS00636; DNAJ_1; 1.
DR PROSITE; PSS0076; DNAJ_2; 1.
KW Lipoprotein.
FT DOMAIN 13 82 DNAJ-LIKE.
FT DOMAIN 118 128 POLY-CYS.
SQ SEQUENCE 198 AA; 22100 MW; 52F98261FBAD978F CRC64;

Query Match 51.3%; Score 40; DB 1; Length 198;
Best Local Similarity 57.1%; Pred. No. 3.8;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AVGHDTKQYRKLA 14
| | : | | | | |
Db 26 ATSDDIRKSYRKLA 39

Search completed: November 8, 2000, 09:03:51
Job time: 861 sec

```
DR PROSITE; PS00092; N6_MTPASE; 1.
KW Transferase; Methyltransferase; Restriction system; Plasmid.
SQ SEQUENCE 269 AA; 30924 MW; 0AB61895696B46EA CRC64;

Query Match 52.6%; Score 41; DB 1; Length 269;
Best Local Similarity 50.0%; Pred. No. 3.5;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VGHDRKQYRKLA 15
Db 245 IGIDSEKYLKIAK 258
:|:|:|:|:|:|
:|:|:|:|:|:|

RESULT 13
SP17_MONDO
ID SP17_MONDO STANDARD; PRT; 179 AA.
AC O62771;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SPERM SURFACE PROTEIN SP17.
CN SP17.
OS Monodelphis domestica (Short-tailed grey opossum).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Monodelphis.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-TESTIS;
RA Wen Y., O'Rand M.G.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SPERM SURFACE ZONA PELLUCIDA BINDING PROTEIN. HELPS TO
CC BIND SPERMATOZOEA TO THE ZONA PELLUCIDA WITH HIGH AFFINITY. MIGHT
CC FUNCTION IN BINDING ZONA PELLUCIDA AND CARBOHYDRATES (BY
CC SIMILARITY).
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: MEMBRANE ASSOCIATED (POTENTIAL).
CC -!- TISSUE SPECIFICITY: TESTIS- AND SPERM-SPECIFIC (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF054290; AAC08025.1; -.
DR INTERPRO; IPR000048; -.
DR PFAM; PF00612; IQ; 1.
KW Membrane.
FT DOMAIN 124 132 POLY-GLU.
SQ SEQUENCE 179 AA; 20864 MW; CADDG13CEC66A00E CRC64;

Query Match 51.3%; Score 40; DB 1; Length 179;
Best Local Similarity 70.0%; Pred. No. 3.4;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GHDRKQYRK 12
Db 156 CHKTRKYLK 165
|||:|:|
|||:|:|

RESULT 14
CSP_BOVIN
ID CSP_BOVIN STANDARD; PRT; 198 AA.
AC Q29455; Q29456;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CYSTEINE STRING PROTEIN (CSP).

OS Mus musculus (Mouse), and Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=MOUSE; TISSUE=BRAIN;
RA Qin N., Lin T., Birnbaumer L.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=RAT; STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
```

```
GN CSP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-ADRENAL MEDULLA;
RX MEDLINE; 96205903.
RA Chamberlain L.H., Burgoyne R.D.;
RT "Identification of a novel cysteine string protein variant and
RT expression of cysteine string proteins in non-neuronal cells.";
RL J. Biol. Chem. 271:7320-7323(1995).
CC -!- FUNCTION: MAY HAVE AN IMPORTANT ROLE IN PRESYNAPTIC FUNCTION. MAY
CC BE INVOLVED IN CALCIUM-DEPENDENT NEUROTRANSMITTER RELEASE AT NERVE
CC ENDINGS.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; CSP1 (SHOWN HERE) AND CSP2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- PTM: FATTY ACYLATED (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS A DNAB-LIKE DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X92666; CAA63354.1; -.
DR EMBL; X92667; CAA63355.1; -.
DR HSPP; P25685; 1HJDJ.
DR INTERPRO; IPR001623; -.
DR PFAM; PF00226; DNAB_1; 1.
DR PROSITE; PS00636; DNAB_1; 1.
DR PROSITE; PS50076; DNAB_2; 1.
KW Lipoprotein; Alternative splicing.
FT DOMAIN 13 82
FT DOMAIN 118 128 DNAB-LIKE.
FT VARSPLIC 165 198 POLY-CYS.
FT EAADTPVIQPASATETITQLTADSHPSYHTDGFN -> GGH
SQ SEQUENCE 198 AA; 22133 MW; 9A3D139FF5428A27 CRC64;

Query Match 51.3%; Score 40; DB 1; Length 198;
Best Local Similarity 57.1%; Pred. No. 3.8;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AVGHDRKQYRKLA 14
Db 26 ATSDDIKKSYRKLA 39
|:|:|:|:|
|:|:|:|:|

RESULT 15
CSP_MOUSE
ID CSP_MOUSE STANDARD; PRT; 198 AA.
AC P54101;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CYSTEINE STRING PROTEIN (CSP).
CN CSP.
OS Mus musculus (Mouse), and Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=MOUSE; TISSUE=BRAIN;
RA Qin N., Lin T., Birnbaumer L.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=RAT; STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
```

```

CC CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC CC or send an email to license@isb-sib.ch).
CC CC
CC EMBL: M12715; AAA26633.1; -.
CC DR PIR: A24545; A24545
CC DR INTERPRO: IPR000734; -.
CC DR PROSITE; PS00120; LIPASE_SER; 1.
CC KW Hydrolase; Lipid degradation; Zymogen; Signal.
CC FT SIGNAL 1 37 POTENTIAL.
CC FT PROPEP 38 295
CC FT CHAIN 296 690 LIPASE.
CC FT DOMAIN 311 690 HYDROPHOBIC.
CC FT ACT_SITE 412 412 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 645 645 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC SQ SEQUENCE 690 AA; 76388 MW; 8E8E3654D0E01A3B CRC64;

Query Match 55.1%; Score 43; DB 1; Length 690;
Best Local Similarity 54.5%; Pred. No. 4.3;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 VGHDRKQYRK 12
DB 591 IGHDAAREWNK 601
:||||:|:|

RESULT 11
IEFS.HUMAN
ID IEFS.HUMAN STANDARD; PRT; 543 AA.
AC P31948;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE TRANSFORMATION-SENSITIVE PROTEIN IEF SSP 3521.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE; 92235077.
RA Honore B., Leffers H., Madsen P., Rasmussen H.H., Vandekerckhove J.,
RA Celis J.E.;
RT "Molecular cloning and expression of a transformation-sensitive human
RT protein containing the TPR motif and sharing identity to the stress-
RT inducible yeast protein STII.";
RL J. Biol. Chem. 267:8485-8491(1992).
RN [2]
RN SEQUENCE OF 101-109; 352-364 AND 374-381.
RP TISSUE-KERATINOCYTES;
RC MEDLINE; 93162043.
RX Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
RA Vandekerckhove J.;
RT "Microsequences of 145 proteins recorded in the two-dimensional gel
RT protein database of normal human epidermal keratinocytes.";
RL Electrophoresis 13:960-969(1992).
CC -!- SIMILARITY: TO YEAST STII.
CC -!- SIMILARITY: CONTAINS 5 TPR DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC CC or send an email to license@isb-sib.ch).
CC CC

```

```

CC DR EMBL; M77184; AAA41811.1; -.
CC DR GCRDB; L19475; AAA68098.1; -.
CC DR GCRDB; GCR_0206; -.
CC DR GCRDB; GCR_0938; -.
CC DR INTERPRO; IPR000832; -.
CC DR INTERPRO; IPR002170; -.
CC DR PFAM; PF00002; 7tm_2; 1.
CC DR PRINTS; PR00249; GPCRSECRETIN.
CC DR PROSITE; PS00649; G_PROTEIN_RECP_F2_1; 1.
CC DR PROSITE; PS00650; G_PROTEIN_RECP_F2_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 26
FT CHAIN 27 591
FT DOMAIN 27 188
FT TRANSMEM 189 212
FT DOMAIN 214 219
FT TRANSMEM 220 239
FT DOMAIN 240 282
FT TRANSMEM 283 306
FT DOMAIN 307 320
FT TRANSMEM 321 342
FT DOMAIN 343 361
FT TRANSMEM 362 382
FT DOMAIN 383 409
FT TRANSMEM 410 428
FT DOMAIN 429 440
FT TRANSMEM 441 463
FT DOMAIN 464 591
FT CARBOHYD 151 151
FT CARBOHYD 161 161
FT CARBOHYD 166 166
FT CARBOHYD 176 176
SQ SEQUENCE 591 AA; 66260 MW; 21944F3051B9E9C1 CRC64;

Query Match 56.4%; Score 44; DB 1; Length 591;
Best Local Similarity 72.7%; Pred. No. 2.4;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRQYRKRLAK 15
Db 398 DTRQYRKLLR 408
|||||:

RESULT 9
VIPR_MELGA
ID VIPR_MELGA STANDARD; PRT; 260 AA.
AC Q91085;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE VASOACTIVE INTESTINAL POLYPEPTIDE RECEPTOR (VIP-R) (VIP RECEPTOR) (FRAGMENT).
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;
RX MEDLINE; 96206340.
RA Xu M., Proudman J.A., Pitts G.R., Wong E.A., Foster D.N.,
RA el Halawani M.E.;
RT "Vasoactive intestinal peptide stimulates prolactin mRNA expression
RT in turkey pituitary cells: effects of dopaminergic drugs.";
RL Proc. Soc. Exp. Biol. Med. 212:52-62(1996).
CC -!- FUNCTION: THIS IS A RECEPTOR FOR VIP. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYL
CC CYCLASE.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.

```

```

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; U31991; AAA99740.1; -.
CC GCRDB; GCR_1749; -.
CC INTERPRO; IPR000832; -.
CC PFAM; PF00002; 7tm_2; 1.
CC PROSITE; PS00649; G_PROTEIN_RECP_F2_1; PARTIAL.
CC PROSITE; PS00650; G_PROTEIN_RECP_F2_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT NON_TER 1 1
FT DOMAIN <1 18 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 19 42 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 43 56 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 57 78 4 (POTENTIAL).
FT DOMAIN 79 95 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 96 119 5 (POTENTIAL).
FT DOMAIN 120 144 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 145 164 6 (POTENTIAL).
FT DOMAIN 165 176 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 177 196 7 (POTENTIAL).
FT DOMAIN 197 260 CYTOPLASMIC (POTENTIAL).
FT SEQUENCE 260 AA; 30358 MW; 15761E6AB5B23D5A CRC64;

Query Match 55.1%; Score 43; DB 1; Length 260;
Best Local Similarity 57.1%; Pred. No. 1.5;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 VGHTRKQYRKLA 15
Db 131 VGHNETSQYSRLAK 144
|||||:

RESULT 10
LIP_STAAU
ID LIP_STAAU STANDARD; PRT; 690 AA.
AC P10335;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE LIPASE PRECURSOR (EC 3.1.1.3) (GLYCEROL ESTER HYDROLASE).
GN GEH.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 86195821.
RA Lee C.Y., Iandolo J.J.;
RT "Lysogenic conversion of staphylococcal lipase is caused by insertion
RT of the bacteriophage L54a genome into the lipase structural gene.";
RL J. Bacteriol. 166:385-391(1986).
RN [2]
RP PROCESSING, AND SEQUENCE OF 296-307.
RC STRAIN=TEN 5;
RX MEDLINE; 92193269.
RA Roloff J., Normark S.;
RT "In vivo processing of Staphylococcus aureus lipase.";
RL J. Bacteriol. 174:1844-1847(1992).
CC -!- CATALYTIC ACTIVITY: TRIACYLGLYCEROL + H(2)O = DIACYLGLYCEROL +
CC A FATTY ACID ANION.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- MISCELLANEOUS: THE EXPRESSION OF STAPHYLOCOCCUS LIPASE IS
CC NEGATIVELY REGULATED BY BACTERIOPHAGE LYSOGENIZATION (LIPASE
CC CONVERSION).
CC -!- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,

```

30-MAY-2000 (Rel. 39, Last annotation update)
 DE PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR
 DE PRECURSOR (PTH/PTHR RECEPTOR).
 GN PTHR1 OR PTHR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=C3H/HEHA;
 RX MEDLINE: 95034305.
 RA Karperien M., van Dijk T.B., Hoeijmakers T., Cremers F.,
 RA Abou-Samra A.-B., Boonstra J., de Laat S.W., Defize L.H.K.;
 RT "Expression pattern of parathyroid hormone/parathyroid hormone
 RT related peptide receptor mRNA in mouse postimplantation embryos
 RT indicates involvement in multiple developmental processes.";
 RL Mech. Dev. 47:29-42(1994).
 RN [2]
 RC SEQUENCE FROM N.A.
 RC STRAIN=BALB/C;
 RX MEDLINE: 94255468.
 RA McCuaig K.A., Clarke J.C., White J.H.;
 RT "Molecular cloning of the gene encoding the mouse parathyroid
 RT hormone/parathyroid hormone-related peptide receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:5051-5055(1994).
 CC -1- FUNCTION: THIS IS A RECEPTOR FOR PARATHYROID HORMONE AND FOR
 CC PARATHYROID HORMONE-RELATED PEPTIDE. THE ACTIVITY OF THIS
 CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATES ADENYLYL
 CC CYCLASE AND ALSO A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER
 CC SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X78936; CAA55536.1; -;
 DR EMBL: L34611; AAA40011.1; -;
 DR EMBL: L34608; AAA40011.1; JOINED.
 DR EMBL: L34607; AAA40011.1; JOINED.
 DR EMBL: L34609; AAA40011.1; JOINED.
 DR EMBL: L34610; AAA40011.1; JOINED.
 DR GCRDB: GCR_1005; -;
 DR GCRDB: GCR_1614; -;
 DR MGD: MGI:97801; PTHR.
 DR INTERPRO: IPR000832; -;
 DR INTERPRO: IPR002170; -;
 DR PFAM: PF00002; 7tmL2; 1.
 DR PRINTS: PR00249; GPCRSECRETIN.
 DR PRINTS: PR00393; PTRHORMONER.
 DR PROSITE: PS00649; G_PROTEIN_RECEPTOR_F2_1; 1.
 DR PROSITE: PS00650; G_PROTEIN_RECEPTOR_F2_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 25
 FT CHAIN 27 591
 FT PARATHYROID HORMONE/PARATHYROID HORMONE-
 FT RELATED PEPTIDE RECEPTOR.
 FT EXTRACELLULAR (POTENTIAL).
 FT 1 (POTENTIAL).
 FT CYTOPLASMIC (POTENTIAL).
 FT 2 (POTENTIAL).
 FT EXTRACELLULAR (POTENTIAL).
 FT 3 (POTENTIAL).
 FT CYTOPLASMIC (POTENTIAL).
 FT 4 (POTENTIAL).
 FT EXTRACELLULAR (POTENTIAL).
 FT 5 (POTENTIAL).
 FT CYTOPLASMIC (POTENTIAL).
 FT 6 (POTENTIAL).

FT DOMAIN 429 440
 FT TRANSSEM 441 463
 FT DOMAIN 464 591
 FT CARBOHYD 151 151
 FT CARBOHYD 161 161
 FT CARBOHYD 166 166
 FT CARBOHYD 176 176
 FT CONFLICT 27 28
 FT CONFLICT 464 465
 FT CONFLICT 500 501
 FT CONFLICT 591 AA; 66313 MW; F7876F8D38BDDDFD CRC64;
 SQ SEQUENCE 591 AA; 66313 MW; F7876F8D38BDDDFD CRC64;
 Query Match 56.4%; Score 44; DB 1; Length 591;
 Best Local Similarity 72.7%; Pred No. 2.4;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 5 DTRKQYRKLLK 15
 DB 398 DTRKQYRKLLK 408
 III:IIIIII:
 RESULT 8
 PTHR_RAT
 ID PTHR_RAT STANDARD; PRT; 591 AA.
 AC P25961;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR
 DE PRECURSOR (PTH/PTHR RECEPTOR).
 GN PTHR1 OR PTHR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RC SEQUENCE FROM N.A.
 RC TISSUE=BONE;
 RX MEDLINE: 92212903.
 RA Abou-Samra A.-B., Jueppner H., Force T., Freeman M.W., Kong X.-F.,
 RA Kronenberg H.M., Segre G.V.;
 RA Schipani E., Urena P., Richards J., Bonventre J.V., Potts J.T. Jr.,
 RT "Expression cloning of a common receptor for parathyroid hormone and
 RT parathyroid hormone-related peptide from rat osteoblast-like cells: a
 RT single receptor stimulates intracellular accumulation of both cAMP
 RT and inositol trisphosphates and increases intracellular free
 RT calcium.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2732-2736(1992).
 RN [2]
 RC SEQUENCE FROM N.A.
 RX MEDLINE: 94292182.
 RA Pausova Z., Bourdon J., Clayton D., Mattei M.-G., Seidlin M.F.,
 RA Janicic N., Riviere M., Szpirer J., Levan G., Szpirer C.;
 RT "Cloning of a parathyroid hormone/parathyroid hormone-related peptide
 RT receptor (PTHr) cDNA from a rat osteosarcoma (UMR 106) cell line:
 RT chromosomal assignment of the gene in the human, mouse, and rat
 RT genomes.";
 RL Genomics 20:20-26(1994).
 CC -1- FUNCTION: THIS IS A RECEPTOR FOR PARATHYROID HORMONE AND FOR
 CC PARATHYROID HORMONE-RELATED PEPTIDE. THE ACTIVITY OF THIS
 CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATES ADENYLYL
 CC CYCLASE AND ALSO A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER
 CC SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

FT VARIANT 410 410 /FTId-VAR_003582.
FT T -> P (IN MURK JANSEN; CONSISTITUTIVELY
FT ACTIVATED).
FT /FTId-VAR_003583.
FT K -> N (IN REF. 2).
FT CONFLICT 471 471
FT CONFLICT 473 473 S -> C (IN REF. 2).
SQ SEQUENCE 593 AA; 66360 MW; DA1400640A6C7F2B CRC64;

Query Match 60.3%; Score 47; DB 1; Length 593;
Best Local Similarity 81.8%; Pred. No. 0.69;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRQYRKRLAK 15
| | | | | | | | |
Db 398 DTRQYRKLLK 408

RESULT 6
CSP_DROME
ID CSP_DROME STANDARD; PRT; 249 AA.
AC Q03751; 061664; 061665; Q0VNV1;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CYSTEINE-STRING PROTEIN (CSP32/CSP29).
GN CSP.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS CSP1 AND CSP3).
RX MEDLINE; 91286850.
RX Zinsmaier K.E., Hofbauer A., Heimbeck G., Pflugfelder G.O.,
RX Buchner S., Buchner E.;
RT "A cysteine-string protein is expressed in retina and brain of
RT Drosophila."
RL J. Neurogenet. 7:15-29(1990).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS CSP1, CSP2 AND CSP3).
RC STRAIN-BERLIN;
RX MEDLINE; 99015937.
RX Eberle K.K., Zinsmaier K.E., Buchner S., Gruhn M., Jenni M.,
RX Arnold C., Leibold C., Reisch D., Walter N., Hafen E., Hofbauer A.,
RX Pflugfelder G.O., Buchner E.;
RT "Wide distribution of the cysteine string proteins in Drosophila
RT tissues revealed by targeted mutagenesis."
RL Cell Tissue Res. 294:203-217(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS CSP1 AND CSP3).
RC STRAIN-BERKELEY;
RX MEDLINE; 20196006.
RX Adams M.D., Celnikner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RX Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RX George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RX Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RX Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RX Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RX Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RX Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RX Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RX Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RX Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RX Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RX de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RX Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RX Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RX Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RX Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RX Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RX Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RX Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

```

```

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; CSP1/CSP32 (SHOWN HERE), CSP2
CC AND CSP3/CSP29; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: BRIAN AND RETINA. PREDOMINANTLY EXPRESSED IN
CC -!- SYNAPTIC TERMINALS OF PHOTORECEPTORS.
CC -!- PTM: FATY ACYLATED (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS A DNAJ-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M63421; AAA28432.1; -
CC EMBL; M63008; AAA28431.1; -
CC EMBL; AF057167; AAD09428.1; -
CC EMBL; AF057167; AAD09430.1; -
CC EMBL; AF057167; AAD09431.1; -
CC EMBL; AE003597; AAF51816.1; -
CC EMBL; AE003597; AAF51817.1; -
CC HSSP; P235685; 1HDJ.
CC FLYBASE; FBgn0004179; Csp.
CC INTERPRO; IPR001623; -
CC INTERPRO; IPR003095; -
CC PFAM; PF00226; DnaJ; 1
CC PRINTS; PR00825; DNAJPROTEIN.
CC PROSITE; PS00636; DNAJ_1; 1.
CC PROSITE; PS0076; DNAJ_2; 1.
CC Lipoprotein; Alternative splicing.
CC DOMAIN 15 84
CC POLY-CYS.
CC FT DOMAIN 121 131
CC FT VARSPLIC 154 174
CC MISSING (IN ISOFORM CSP3 AND ISOFORM
CC CSP2).
CC VARSPLIC 243 249
CC DMVNOQY -> GI (IN ISOFORM CSP3).
CC VARIANT 71 71
CC N -> D.
CC SEQUENCE 249 AA; 26896 MW; 3EF97C3BF2553EB8 CRC64;

Query Match 57.7%; Score 45; DB 1; Length 249;
Best Local Similarity 64.3%; Pred. No. 0.61;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AVGHDTKQYRKLA 14
| | | | | | | | |
Db 28 ATGDDIKTKYRKLA 41

RESULT 7
PTRR_MOUSE
ID PTRR_MOUSE STANDARD; PRT; 591 AA.
AC P41593; Q62119;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)

```


the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M74445; AAA30979.1; -
 GCRDB; GCR_0204; -
 INTERPRO; IPR000832; -
 INTERPRO; IPR002170; -
 PFAM; PF00002; 7tm_2; 1.
 PRINTS; PR00249; GPCRSECRETIN.
 PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
 PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
 G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 SIGNAL 1 26
 CHAIN 27 585

DOMAIN 27 185
 TRANSMEM 186 209
 DOMAIN 210 216
 TRANSMEM 217 236
 DOMAIN 237 276
 TRANSMEM 277 300
 DOMAIN 301 314
 TRANSMEM 315 336
 DOMAIN 337 355
 TRANSMEM 356 376
 DOMAIN 377 403
 TRANSMEM 404 422
 DOMAIN 423 434
 TRANSMEM 435 457
 DOMAIN 458 585
 CARBOHYD 148 148
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CARBOHYD 158 158
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CARBOHYD 163 163
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CARBOHYD 173 173
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SEQUENCE 585 AA; 65963 MW; 34900384CD6DF477 CRC64;

Query Match 60.3%; Score 47; DB 1; Length 585;
 Best Local Similarity 81.8%; Pred. No. 0.68;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRKQYRKLA 15
 |||:|||||
 Db 392 DTRQYRKLLK 402

RESULT 4
 ID PTRR_PIG STANDARD; PRT; 585 AA.
 AC P50133;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR
 DE PRECURSOR (PTH/PTHr RECEPTOR).
 GN PTHR1 OR PTHR.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 96305358.
 RA Black E.C., Smith D.P., Zhang X.Y., Frolik C.A., Harvey A.,
 RA Chandrasekhar S., Hsiung H.M.;
 RT "Structure and functional expression of a complementary DNA for
 RT porcine parathyroid hormone/parathyroid hormone-related peptide
 RT receptor.";
 RL Biochim. Biophys. Acta 1307:339-347(1996).

-!- FUNCTION: THIS IS A RECEPTOR FOR PARATHYROID HORMONE AND FOR PARATHYROID HORMONE-RELATED PEPTIDE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL CYCLASE AND ALSO A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM (BY SIMILARITY).
 -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U18315; AAC48619.1; -
 GCRDB; GCR_1607; -
 INTERPRO; IPR000832; -
 INTERPRO; IPR002170; -
 PFAM; PF00002; 7tm_2; 1.
 PRINTS; PR00249; GPCRSECRETIN.
 PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
 PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
 G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 SIGNAL 1 26
 CHAIN 27 585

DOMAIN 27 184
 TRANSMEM 185 208
 DOMAIN 209 215
 TRANSMEM 216 235
 DOMAIN 236 277
 TRANSMEM 278 301
 DOMAIN 302 315
 TRANSMEM 316 337
 DOMAIN 338 356
 TRANSMEM 357 377
 DOMAIN 378 404
 TRANSMEM 405 423
 DOMAIN 424 435
 TRANSMEM 436 458
 DOMAIN 459 585
 CARBOHYD 147 147
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CARBOHYD 157 157
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CARBOHYD 162 162
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CARBOHYD 172 172
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SEQUENCE 585 AA; 65682 MW; 60BE15CD49B7D210 CRC64;

Query Match 60.3%; Score 47; DB 1; Length 585;
 Best Local Similarity 81.8%; Pred. No. 0.68;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRKQYRKLA 15
 |||:|||||
 Db 393 DTRQYRKLLK 403

RESULT 5
 ID PTRR_HUMAN STANDARD; PRT; 593 AA.
 AC Q03431;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR
 DE PRECURSOR (PTH/PTHr RECEPTOR).
 GN PTHR1 OR PTHR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: November 8, 2000, 09:03:49 ; Search time 58.45 seconds
(without alignments)
8.199 Million cell updates/sec

Title: US-09-236-468A-2_COPY_349_363
Perfect score: 78
Sequence: 1 AVGHDTKQYRKLA 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 87993 seqs, 31947931 residues
Total number of hits satisfying chosen parameters: 87993

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	550	1	PTR2_HUMAN
2	72	92.3	546	1	PTH2_RAT
3	47	60.3	585	1	PTRR_DIDMA
4	47	60.3	585	1	PTRR_PIG
5	47	60.3	593	1	PTRR_HUMAN
6	45	57.7	249	1	CSP_DROME
7	44	56.4	591	1	PTRR_MOUSE
8	44	56.4	591	1	PTRR_RAT
9	43	55.1	260	1	VIPR_MELGA
10	43	55.1	690	1	LIP_STAAN
11	42	53.8	543	1	IEFS_HUMAN
12	41	52.6	269	1	ML22_LACLC
13	40	51.3	179	1	SP17_MONDO
14	40	51.3	198	1	CSP_BOVIN
15	40	51.3	198	1	CSP_MOUSE
16	40	51.3	258	1	TPIS_STRCO
17	39	50.0	559	1	V269_HUMAN
18	39	50.0	1432	1	WRN_HUMAN
19	38	48.7	222	1	GPA3_RAT
20	38	48.7	437	1	VIPS_MOUSE
21	38	48.7	437	1	VIPS_RAT
22	38	48.7	438	1	VIPS_HUMAN
23	38	48.7	496	1	RECO_BACSU
24	38	48.7	511	1	MDJL_YEAST
25	38	48.7	834	1	PPSA_STAMA
26	38	48.7	910	1	DNJM_MYCPN
27	37	47.4	140	1	RL11_THEMA
28	37	47.4	195	1	CSP_TORCA
29	37	47.4	236	1	RR2_PEA
30	37	47.4	243	1	TPIS_MYCHR
31	37	47.4	268	1	MT22_STRPN
32	37	47.4	333	1	KC21_ARATH
33	37	47.4	333	1	KC22_ARATH

34	37	47.4	355	1	YRY1_CABEL	Q10005 caenorhabdl
35	37	47.4	392	1	YAY1_SCHPO	Q10209 schizosacch
36	37	47.4	417	1	TRMU_YEAST	Q12093 saccharomyc
37	37	47.4	631	1	COBT_PSEDE	Q29934 pseudomonas
38	37	47.4	746	1	PPEA_PSEAE	Q05098 pseudomonas
39	37	47.4	1416	1	BLM_MOUSE	Q08700 mus musculu
40	37	47.4	1417	1	BLM_HUMAN	Q34132 homo sapien
41	37	47.4	1682	1	CIN6_HUMAN	Q01118 homo sapien
42	36	46.2	123	1	WN3A_ALOVU	P28101 aloplas vul
43	36	46.2	222	1	GTA4_MOUSE	P24472 mus musculu
44	36	46.2	241	1	HSJ2_HUMAN	O75190 homo sapien
45	36	46.2	307	1	DNAJ_DEIPR	O34136 delnocooccu

ALIGNMENTS

RESULT 1

PTR2_HUMAN

ID PTR2_HUMAN STANDARD; PRT; 550 AA.

AC P49190;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DE 01-OCT-2000 (Rel. 40, Last annotation update)

DE PARATHYROID HORMONE RECEPTOR PRECURSOR (PTH2 RECEPTOR).

GN PTHR2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

RN [1]

RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.

RC TISSUE=BRAIN;

RX MEDLINE: 95318121.

RA Usdin T.B., Gruber C., Bonner T.I.;

RT "Identification and functional expression of a receptor selectively recognizing parathyroid hormone, the PTH2 receptor.";

RL J. Biol. Chem. 270:15455-15458(1995).

RN [2]

RP SEQUENCE OF 26-40 AND 306-550 FROM N.A.

RX MEDLINE: 97079671.

RA Usdin T.B., Modi W., Bonner T.I.;

RT "Assignment of the human PTH2 receptor gene (PTH2) to chromosome 2q33 by fluorescence in situ hybridization.";

RL Genomics 37:140-141(1996).

CC -!- FUNCTION: THIS IS A SPECIFIC RECEPTOR FOR PARATHYROID HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYL CYCLASE.

CC -!- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN BRAIN AND PANCREAS.

CC ALSO EXPRESSED IN THE TESTIS.

CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; U25128; AAC50157.1; -

DR EMBL; U47124; AAA96796.1; -

DR EMBL; U47129; AAC50767.1; -

DR EMBL; U47125; AAC50767.1; JOINED.

DR EMBL; U47126; AAC50767.1; JOINED.

DR EMBL; U47127; AAC50767.1; JOINED.

DR EMBL; U47128; AAC50767.1; JOINED.

DR GCRDB; GCR.2003; -

DR MIM; 601469; -

DR INTERPRO; IPR000832; -

DR PFAM; PF00002; 7tm_2; 1.

DR PRINTS; PR00249; GPCRSECRETIN.

DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.

DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.

Search completed: November 8, 2000, 08:53:30
Job time: 361 sec

A:Accession: A42698
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-585,'G',587-591 <ABO>
A:Experimental source: R08 17/2.8 osteosarcoma cells
A:Note: sequence extracted from NCBI backbone (NCBIP:92187)
C:Superfamily: glucagon receptor

Query Match 56.4%; Score 44; DB 2; Length 591;
Best Local Similarity 72.7%; Pred. No. 7.4; 1; Indels 0; Gaps 0;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRKQYRKLA 15
| | | | |
Db 398 DTRQYRKLRL 408

RESULT 7
A24545
triacylglycerol lipase (EC 3.1.1.3) - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 08-Oct-1999
C:Accession: A24545
R:Lee, C.Y.; Iandolo, J.J.
A:Title: Lyso-genic conversion of staphylococcal lipase is caused by insertion of the bac
A:Reference number: A24545; MUID:86195821
A:Accession: A24545
A:Molecule type: DNA
A:Residues: 1-690 <LEE>
A:Cross-references: GB:M12715; NID:g153019; PIDN:AAA26633.1; PID:g153020
C:Superfamily: Staphylococcus triacylglycerol lipase
C:Keywords: carboxylic ester hydrolase

Query Match 55.1%; Score 43; DB 2; Length 690;
Best Local Similarity 54.5%; Pred. No. 13; 1; Indels 0; Gaps 0;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VGHTRKQYRK 12
| | | | |
Db 591 IGHDAEEWRK 601

RESULT 8
T16087
hypothetical protein F18C5.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16087
R:Fulton, L.
Submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid F18C5.
A:Reference number: Z18459
A:Accession: T16087
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-986 <FUL>
A:Cross-references: EMBL:U29097; NID:g861394; PID:g861396; PIDN:AAA68410.1; CESP:F18C5.2
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F18C5.2
A:Introns: 39/1; 58/1; 317/2; 341/1; 468/1; 496/3; 524/2; 596/3; 705/2; 738/3; 780/2; 82

Query Match 55.1%; Score 43; DB 2; Length 986;
Best Local Similarity 61.5%; Pred. No. 18; 4; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GHDRKQYRKLA 15
| | | | |
Db 357 GHDFNSYRHLAE 369

RESULT 9
T22493
hypothetical protein F52B5.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22493
R:Kershaw, J.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19570
A:Accession: T22493
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1425 <WIL>
A:Cross-references: EMBL:Z75541; PIDN:CAA99855.1; GSPDB:GN00019; CESP:F52B5.3
A:Experimental source: clone F52B5
C:Genetics:
A:Gene: CESP:F52B5.3
A:Map position: 1
A:Introns: 22/2; 178/2; 269/2; 439/2; 564/3; 663/3; 705/3; 918/1; 1050/2; 1129/1; 116

Query Match 55.1%; Score 43; DB 2; Length 1425;
Best Local Similarity 63.6%; Pred. No. 26; 0; Indels 0; Gaps 0;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 4 HDTRKQYRKLA 14
| | | | |
Db 35 NDTRKRYRRA 45

RESULT 10
T18661
hypothetical protein B0035.14 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18661
R:White, S.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z19002
A:Accession: T18661
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-401 <WIL>
A:Cross-references: EMBL:Z73102; PIDN:CAA97416.1; GSPDB:GN00022; CESP:B0035.14
A:Experimental source: clone B0035
C:Genetics:
A:Gene: CESP:B0035.14
A:Map position: 4
A:Introns: 22/3; 47/1; 180/1; 233/1; 310/2; 365/1

Query Match 53.8%; Score 42; DB 2; Length 401;
Best Local Similarity 80.0%; Pred. No. 11; 1; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRKQYRKLA 14
| | | | |
Db 152 DIRKEYRKLA 161

RESULT 11
A38093
transformation-sensitive protein IEF SSP 3521 - human
C:Species: Homo sapiens (man)
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 11-Jan-2000
C:Accession: A38093
R:Honore, B.; Leffers, H.; Madsen, P.; Rasmussen, H.H.; Vandekerckhove, J.; Celis, J.
J. Biol. Chem. 267, 8485-8491, 1992
A:Title: Molecular cloning and expression of a transformation-sensitive human protein
A:Reference number: A38093; MUID:92235077
A:Accession: A38093

Query Match 60.3%; Score 47; DB 2; Length 585;
Best Local Similarity 81.8%; Pred. No. 2.2;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRQYRKLLK 15
|||||

Db 392 DTRQYRKLLK 402

RESULT 3
A49191
parathyroid hormone/PTH-related peptide receptor - human
N:Alternate names: parathyroid hormone/parathyroid hormone related peptide receptor
C:Species: Homo sapiens (man)
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 11-Jan-2000
R:Schipani, E.; Weinstein, L.S.; Bergwitz, C.; Iida-Klein, A.; Kong, X.F.; Stuhmann, M.
Kronenberg, H.M.; Abou-Samra, A.B.; Segre, G.V.; Juppner, H.
J. Clin. Endocrinol. Metab. 80, 1611-1621, 1995
A:Title: Pseudohypoparathyroidism type 1b is not caused by mutations in the coding exons
A:Reference number: 138139; MUID:95263723
A:Accession: 138139
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-593 <RES>
A:Cross-references: EMBL:U22409; NID:g897594; PIDN:AA60657.1; PID:g897596
R:Schipani, E.; Karga, H.; Karaplis, A.C.; Potts Jr., J.T.; Kronenberg, H.M.; Segre, G.V.
Endocrinology 132, 2157-2165, 1993
A:Title: Identical complementary deoxyribonucleic acids encode a human renal and bone pa
A:Reference number: A49191; MUID:93238641
A:Accession: A49191
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-593 <SCH>
A:Cross-references: GB:L04308; NID:q150721; PIDN:AAA36525.1; PID:q150722
A:Note: sequence extracted from NCBI backbone (NCBIN:130233, NCBI:130234)
R:Schneider, H.; Feyen, J.H.; Seuwen, K.; Movva, N.R.
Eur. J. Pharmacol. 246, 149-155, 1993
A:Title: Cloning and functional expression of a human parathyroid hormone receptor.
A:Reference number: 138113; MUID:93387403
A:Accession: 138113
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-593 <RE2>
A:Cross-references: EMBL:X68596; NID:g396812; PIDN:CAA48589.1; PID:g396813
R:Levine, M.
submitted to the EMBL Data Library, November 1994
A:Reference number: G07787
A:Accession: G01562
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-593 <LEV>
A:Cross-references: EMBL:U17418; NID:g956129; PIDN:AAA56774.1; PID:g956130
C:Genetics;
A:Introns: 25/3; 60/1; 105/1; 142/1; 181/3; 213/2; 278/3; 330/1; 350/2; 372/3; 404/2; 45
C:Superfamily: glucagon receptor
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 60.3%; Score 47; DB 2; Length 593;
Best Local Similarity 81.8%; Pred. No. 2.2;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRQYRKLLK 15
|||||

Db 398 DTRQYRKLLK 408

RESULT 4
159297
parathyroid hormone/parathyroid hormone related-peptide receptor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 11-Jan-2000
A:Reference number: A42698; MUID:92212903

C:Accession: I59297
R:McCuag, K.A.; Clarke, J.C.; White, J.H.
Proc. Natl. Acad. Sci. U.S.A. 91, 5051-5055, 1994
A:Title: Molecular cloning of the gene encoding the mouse parathyroid hormone/parathy
A:Reference number: 159297; MUID:94255468
A:Accession: I59297
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-589 <RES>
A:Cross-references: GB:L34611; NID:g530149; PIDN:AAA40011.1; PID:g530151
C:Genetics;
A:Gene: PTHR
A:Introns: 25/3; 60/1; 105/1; 142/1; 181/3; 213/2; 278/3; 330/1; 350/2; 372/3; 404/2;
C:Superfamily: glucagon receptor

Query Match 56.4%; Score 44; DB 2; Length 589;
Best Local Similarity 72.7%; Pred. No. 7.3;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRQYRKLLK 15
|||||

Db 398 DTRQYRKLLR 408

RESULT 5
S44203
parathyroid hormone-related peptide receptor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 11-Jan-2000
C:Accession: S44203
R:Karperien, M.; van Dijk, T.B.; Hoeijmakers, T.; Cremers, F.; Abou-Samra, A.B.; Boon
submitted to the EMBL Data Library, April 1994
A:Description: Expression pattern of parathyroid hormone/parathyroid hormone related
A:Reference number: S44203
A:Accession: S44203
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-591 <KAR>
A:Cross-references: EMBL:X78936; NID:g474828; PIDN:CAA55536.1; PID:g474829
C:Superfamily: glucagon receptor

Query Match 56.4%; Score 44; DB 2; Length 591;
Best Local Similarity 72.7%; Pred. No. 7.4;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRQYRKLLK 15
|||||

Db 398 DTRQYRKLLR 408

RESULT 6
154195
parathyroid hormone/parathyroid hormone related-peptide receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 11-Jan-2000
C:Accession: 154195; A42698
R:Pausova, Z.; Bourdon, J.; Clayton, D.; Mattei, M.G.; Seldin, M.F.; Janicic, N.; Riv
Genomics 20, 20-26, 1994
A:Title: Cloning of a parathyroid hormone/parathyroid hormone-related peptide recepto
and rat genomes.
A:Reference number: 154195; MUID:94292182
A:Accession: 154195
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-591 <RES>
A:Cross-references: GB:L19475; NID:g467316; PIDN:AAA68098.1; PID:g467317
R:Abou-Samra, A.B.; Juppner, H.; Force, T.; Freeman, M.W.; Kong, X.F.; Schipani, E.;
Proc. Natl. Acad. Sci. U.S.A. 89, 2732-2736, 1992
A:Title: Expression cloning of a common receptor for parathyroid hormone and parathyri
n of both cAMP and inositol trisphosphates and increases intracellular free calcium.
A:Reference number: A42698; MUID:92212903

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:53:28 ; Search time 99.87 Seconds
(without alignments)
9.531 Million cell updates/sec

Title: US-09-236-468A-2_COPY_349_363

Perfect score: 78

Sequence: 1 AVGHDTTRKQYRKLA 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 182106

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_65:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	550	2 A57519	parathyroid hormon
2	47	60.3	585	2 A39286	parathyroid hormon
3	47	60.3	593	2 A49191	parathyroid hormon
4	44	56.4	589	2 I59297	parathyroid hormon
5	44	56.4	591	2 S44203	parathyroid hormon
6	44	56.4	591	2 I54195	parathyroid hormon
7	43	55.1	690	2 A24545	triacylglycerol li
8	43	55.1	986	2 T16087	hypothetical prote
9	43	55.1	1425	2 T22493	hypothetical prote
10	42	53.8	401	2 T18661	hypothetical prote
11	42	53.8	543	2 A38093	transformation-sen
12	40	51.3	167	2 S70516	cysteine string pr
13	40	51.3	198	2 I52655	cysteine string pr
14	40	51.3	198	2 S70515	cysteine string pr
15	40	51.3	258	2 T36018	triose phosphate 1
16	40	51.3	719	2 T35189	probable ATP-depen
17	40	51.3	2241	2 T02857	conserved hypothet
18	39	50.0	320	2 T03899	hypothetical prote
19	38	48.7	150	2 T33554	hypothetical prote
20	38	48.7	222	1 XURT9C	glutathione transf
21	38	48.7	419	2 S42989	T48 protein - frui
22	38	48.7	437	2 JU0185	PACAP/VIP receptor
23	38	48.7	437	2 S39069	vasoactive intesti
24	38	48.7	438	2 G02822	vasoactive intesti
25	38	48.7	479	2 T30997	hypothetical prote
26	38	48.7	496	2 A96631	ATP-dependent DNA
27	38	48.7	511	2 S38898	heat shock protein
28	38	48.7	824	2 G75413	DNA helicase RecQ
29	38	48.7	834	2 S51006	pyruvate,water dik

30	38	48.7	910	1 S73361	dnaj homolog prote
31	38	48.7	1231	2 T24415	hypothetical prote
32	38	48.7	1436	2 T14895	DNA helicase 1 - A
33	38	48.7	2062	2 A59297	myosin X - mouse
34	37	47.4	141	1 RSHG11	ribosomal protein
35	37	47.4	195	2 JH0719	omega-conotoxin re
36	37	47.4	200	2 C81403	hypothetical prote
37	37	47.4	236	1 R3PM2	ribosomal protein
38	37	47.4	256	2 A24372	dnaj protein - Str
39	37	47.4	256	2 A81370	probable dnaj-like
40	37	47.4	328	2 G71610	protein with DnaJ
41	37	47.4	333	2 S31098	casein kinase II (
42	37	47.4	333	2 S31099	casein kinase II (
43	37	47.4	333	2 T45853	CASEIN KINASE II,
44	37	47.4	338	2 H69022	hypothetical prote
45	37	47.4	355	2 T24938	hypothetical prote

ALIGNMENTS

RESULT 1

A57519

parathyroid hormone receptor 2 precursor - human

N:Alternate names: PTH2 receptor

C:Species: Homo sapiens (man)

C>Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 24-Nov-1999

C:Accession: A57519

R:Usdin, T.B.; Gruber, C.; Bonner, T.I.

J. Biol. Chem. 270, 15455-15458, 1995

A:Title: Identification and functional expression of a receptor selectively recognizi

A:Reference number: A57519; MUID:95318121

A:Accession: A57519

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-550 <US>

A:Cross-references: GB:U25128; MID:g887966; PIDN:AAC50157.1; PID:g887967

C:Genetics:

A:Gene: GDB:PTHR2; PTHR2R

A:Cross-references: GDB:731977; OMIM:601469

A:Map position: 2q33-2q33

C:Superfamily: glucagon receptor

C:Keywords: hormone receptor

Query Match Best Local Similarity 100.0%; Score 78; DB 2; Length 550;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVGHDTTRKQYRKLA 15

DB 349 AVGHDTTRKQYRKLA 363

RESULT 2

A39286

parathyroid hormone / parathyroid hormone-related peptide - North American opossum

C:Species: Didelphis virginiana, Didelphis marsupialis virginiana (North American opo

C>Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 11-Jan-2000

C:Accession: A39286

R:Jueppner, H.; Abou-Samra, A.B.; Freeman, M.; Kong, X.F.; Schipani, E.; Richards, J.

Science 254, 1024-1026, 1991

A:Title: A G protein-linked receptor for parathyroid hormone and parathyroid hormone-

A:Reference number: A39286; MUID:92054592

A:Accession: A39286

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-585 <JUP>

A:Cross-references: GB:M74445

C:Superfamily: glucagon receptor

C:Keywords: G protein-coupled receptor; transmembrane protein

THIS PAGE BLANK (USPTO)

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 52 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-293-284A-10

Query Match 57.7%; Score 45; DB 2; Length 52;
Best Local Similarity 64.3%; Pred. No. 0.43;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AVGHDTKQYRKLA 14
| | | | |
Db 1 ATGDDIKTKYRKLA 14

RESULT 15
US-08-468-249A-20
; Sequence 20, Application US/08468249A
; Patent No. 5886148
; GENERAL INFORMATION:
; APPLICANT: Segre et al., Gino V.
; TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR AND DNA
; TITLE OF INVENTION: ENCODING SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,249A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,475
; FILING DATE: 06-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,702
; FILING DATE: 04-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/071003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-249A-20

Query Match 56.4%; Score 44; DB 2; Length 591;
Best Local Similarity 72.7%; Pred. No. 6.6;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRKQYRKLA 15

Db 398 DTRQYRKLLR 408
| | | | | :
Search completed: November 8, 2000, 08:49:19
Job time: 112 sec

GENERAL INFORMATION:
APPLICANT: Segre et al., Gino V.
TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR AND DNA
TITLE OF INVENTION: ENCODING SAME
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,249A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,475
FILING DATE: 06-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/681,702
FILING DATE: 04-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00786/071003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 593 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-468-249A-21

Query Match 60.3%; Score 47; DB 2; Length 593;
Best Local Similarity 81.8%; Pred. No. 2.1;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 DTRKQYRKLA 15
||||:||||| |
Db 398 DTRQYRKLLK 408

RESULT 13
US-08-346-849-10
Sequence 10, Application US/08346849
Patent No. 5670483
GENERAL INFORMATION:
APPLICANT: Zhang, Shuguang
APPLICANT: Lockshin, Curtis
APPLICANT: Rich, Alexander
TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173-4799
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/346,849
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,326
FILING DATE: 28 DECEMBER 1992
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-6008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 52 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-346-849-10

Query Match 57.7%; Score 45; DB 1; Length 52;
Best Local Similarity 64.3%; Pred. No. 0.43;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AVGHDKRKQYRKLA 14
| | | : | | | | |
Db 1 ATGDDIKKTKYRKLA 14

RESULT 14
US-08-293-284A-10
Sequence 10, Application US/08293284A
Patent No. 5955343
GENERAL INFORMATION:
APPLICANT: Holmes, Todd
APPLICANT: Zhang, Shuguang
APPLICANT: Rich, Alexander
APPLICANT: DiPersio, C. Michael
APPLICANT: Lockshin, Curtis
TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/293,284A
FILING DATE: 22-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,326
FILING DATE: 28-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-6008A

```
; TELEFAX: (415) 854-8275
; INFORMATION FOR SEQ ID NO: 125:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1-585
; OTHER INFORMATION: /note= "PTH receptor"
;
US-08-142-551B-125
```

```
Query Match 60.3% Score 47: DB 2; Length 585;
Best Local Similarity 81.8%; Pred. No. 2.1;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 5 DTRQYRKRLK 15
   |||:|||||
Db 392 DTRQYRKLLK 402
```

```
RESULT 10
US-08-869-477-6
; Sequence 6, Application US/08869477
; Patent No. 5846747
; GENERAL INFORMATION:
; APPLICANT: Thorens, Bernard
; TITLE OF INVENTION: Receptor for the Glucagon-Like-Peptide-1
; TITLE OF INVENTION: (GLP-1)
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58467470 No. 5846747disk of No. 5846747th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/869,477
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,439
; FILING DATE: 24-NOV-93
; APPLICATION NUMBER: DK 398/92
; FILING DATE: 25-MAR-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PC/EP93/00697
; FILING DATE: 23-MAR-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Harrington, James J.
; REGISTRATION NUMBER: 38,711
; REFERENCE/DOCKET NUMBER: 3756.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
```

```
; ORGANISM: Didelphis virginiana
; US-08-869-477-6
;
Query Match 60.3% Score 47: DB 2; Length 585;
Best Local Similarity 81.8%; Pred. No. 2.1;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRQYRKRLK 15
   |||:|||||
Db 392 DTRQYRKLLK 402

RESULT 11
US-08-468-249A-19
; Sequence 19, Application US/08468249A
; Patent No. 5886148
; GENERAL INFORMATION:
; APPLICANT: Segre et al., Gino V.
; TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR AND DNA
; TITLE OF INVENTION: ENCODING SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,249A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,475
; FILING DATE: 06-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,702
; FILING DATE: 04-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/071003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-249A-19
```

```
Query Match 60.3% Score 47: DB 2; Length 585;
Best Local Similarity 81.8%; Pred. No. 2.1;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 5 DTRQYRKRLK 15
   |||:|||||
Db 392 DTRQYRKLLK 402
```

```
RESULT 12
US-08-468-249A-21
; Sequence 21, Application US/08468249A
; Patent No. 5886148
```

; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,249A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,475
; FILING DATE: 06-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,702
; FILING DATE: 04-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/071003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 515 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-249A-18

Query Match 60.3%; Score 47; DB 2; Length 515;
Best Local Similarity 81.8%; Pred. No. 1.8;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 DTRKOYRKLLK 15
|111111111|
Db 392 DTRQYRKLLK 402

RESULT 8
US-08-142-439A-6
; Sequence 6, Application US/08142439A
; Patent No. 5670360
; GENERAL INFORMATION:
; APPLICANT: Thorens, Bernard
; TITLE OF INVENTION: Receptor for the Glucagon-Like-Peptide-1
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5670360o No. 5670360disk of No. 5670360th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,439A
; FILING DATE: 24-NOV-93
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 398/92
; FILING DATE: 25-MAR-92
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP93/00697
; FILING DATE: 23-MAR-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Harrington, James J.
; REGISTRATION NUMBER: 38,711
; REFERENCE/DOCKET NUMBER: 3756.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Didelphis virginiana
; US-08-142-439A-6

Query Match 60.3%; Score 47; DB 1; Length 585;
Best Local Similarity 81.8%; Pred. No. 2.1;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 DTRKOYRKLLK 15
|111111111|
Db 392 DTRQYRKLLK 402

RESULT 9
US-08-142-551B-125
; Sequence 125, Application US/08142551B
; Patent No. 5814603
; GENERAL INFORMATION:
; APPLICANT: Oldenburg, Kevin R.
; TITLE OF INVENTION: COMPOUNDS WITH PTH ACTIVITY AND
; TITLE OF INVENTION: RECOMBINANT DNA VECTORS ENCODING SAME
; NUMBER OF SEQUENCES: 132
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: US
; ZIP: 22313
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,551B
; FILING DATE: 25-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/077,296
; FILING DATE: 14-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/898,219
; FILING DATE: 12-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/965,677
; FILING DATE: 22-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Swiss, Gerald F.
; REGISTRATION NUMBER: 30,113
; REFERENCE/DOCKET NUMBER: 000324-010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-7400

; REFERENCE/DOCKET NUMBER: 00786/071002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-07-864-475A-13

Query Match 60.3%; Score 47; DB 1; Length 25;
Best Local Similarity 81.8%; Pred. No. 0.098;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRQYRKLLK 15
Db 15 DTRQYRKLLK 25

RESULT 5
US-08-468-249A-13
; Sequence 13, Application US/08468249A
; Patent No. 5886148
; GENERAL INFORMATION:
; APPLICANT: Segre et al., Gino V.
; TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR AND DNA
; TITLE OF INVENTION: ENCODING SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,249A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER: US 07/864,475
; FILING DATE: 06-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,702
; FILING DATE: 04-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/071003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-468-249A-13

Query Match 60.3%; Score 47; DB 2; Length 25;
Best Local Similarity 81.8%; Pred. No. 0.098;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 5 DTRQYRKLLK 15
Db 15 DTRQYRKLLK 25

RESULT 6
US-08-468-011A-14
; Sequence 14, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Oistein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-011A-14

Query Match 60.3%; Score 47; DB 3; Length 52;
Best Local Similarity 81.8%; Pred. No. 0.2;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRQYRKLLK 15
Db 20 DTRQYRKLLK 30

RESULT 7
US-08-468-249A-18
; Sequence 18, Application US/08468249A
; Patent No. 5886148
; GENERAL INFORMATION:
; APPLICANT: Segre et al., Gino V.
; TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR AND DNA
; TITLE OF INVENTION: ENCODING SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street

Db 16 AVGHDRKQYRKLAK 30

RESULT 2

US-08-468-011A-2
; Sequence 2, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (P201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-011A-2

Query Match 100.0%; Score 78; DB 3; Length 541;
Best Local similarity 100.0%; Pred. No. 1.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVGHDRKQYRKLAK 15

Db 349 AVGHDRKQYRKLAK 363

RESULT 3

PCT-US95-07085-2
; Sequence 2, Application PC/TUS9507085
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone Receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland

; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07085
; FILING DATE: 05-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-393
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-07085-2

Query Match 100.0%; Score 78; DB 4; Length 541;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVGHDRKQYRKLAK 15

Db 349 AVGHDRKQYRKLAK 363

RESULT 4

US-07-864-475A-13
; Sequence 13, Application US/07864475A
; Patent No. 5494806
; GENERAL INFORMATION:
; APPLICANT: Segre, Gino V.
; APPLICANT: Kronenberg, Henry M.
; APPLICANT: Abou-Samra, Abdul-Badi
; APPLICANT: Juppner, Harald
; APPLICANT: Potts, John T. [Jr.]
; APPLICANT: Schipani, Ernestina
; TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR
; TITLE OF INVENTION: AND DNA ENCODING SAME
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/864,475A
; FILING DATE: 04-06-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/681,702
; FILING DATE: 05-04-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162

PN US5494806-A.
 XX
 PD 27-FEB-1996.
 XX
 PF 05-APR-1991; 91US-0681702.
 XX
 PR 06-APR-1992; 92US-0864475.
 PR 05-APR-1991; 91US-0681702.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;
 XX
 DR N-PSDB; T15947.
 DR
 XX
 PT DNA encoding vertebrate parathyroid hormone receptor - useful for
 PT diagnosis and treatment of e.g. hypercalcaemia, hypo-calcaemia,
 PT cancer etc.
 XX
 PS Claim 1; Fig 3A-3E; 64pp; English.
 XX
 CC A rat parathyroid hormone/parathyroid hormone-related protein
 CC (PTHrP/PTHrP) receptor (R92277) is encoded by cDNA clone R15B
 CC (T15947) isolated from a rat osteosarcoma ROS 17/2.8 cDNA library.
 CC The receptor is a G-protein linked receptor having 7 transmembrane
 CC domains. It induces an increase in intracellular cAMP and calcium
 CC upon challenge with PTH or PTHrP. Recombinant receptor can be
 CC produced in vector/host cell systems and used in the treatment,
 CC diagnosis or prognosis of e.g. hypercalcaemia and hypocalcaemia,
 CC to screen for (ant)agonists and to raise antibodies. Host cells
 CC expressing the receptor are used for diagnostic measurement of PTH
 CC serum levels.
 XX
 SQ Sequence 591 AA;

Query Match 56.4%; Score 44; DB 17; Length 591;
 Best Local Similarity 72.7%; Pred. No. 5.9;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRKQYRKLA 15
 |||:|||||:
 Db 398 dtrgqyrklr 408

Search completed: November 8, 2000, 08:51:45
 Job time: 257 sec

PT - for (differential) diagnosis of hypercalcaemia, and diagnosis
 XX and treatment of tumours
 PS Claim 22; Fig 6; 91pp; English.
 XX
 CC The protein sequence was deduced from the cDNA sequence obtd by
 CC screening a human kidney oligo dr-primed cDNA library in lambda
 CC 9c10 and a genomic library of human placental DNA in EMBL3 with a
 CC probe comprising most of the coding sequence of rat bone parathyroid
 CC hormone/parathyroid hormone related protein (PTH/PTHrP) receptor
 CC protein. The clone encodes a protein which may be used in a
 CC therapeutic compsn. to inhibit activation of PTH or PTHrP and thus
 CC reduce the level of calcium in the blood. Cpbs. capable of competing
 CC with PTH or PTHrP for binding can be identified using the protein prod.
 CC and DNAs homologous to PTH DNA can be identified using fragments of the
 CC clone as probes. The sequence may be used for the prodn. of antibodies
 CC useful for the treatment, classification, prognosis and/or treatment of
 CC disorders related to the interaction between a cell receptor and a
 CC ligand such as in hypercalcaemia. See also R27704-16.
 XX
 SQ Sequence 614 AA;
 Query Match 60.3%; Score 47; DB 13; Length 614;
 Best Local Similarity 81.8%; Pred. No. 1.8;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 5 DTRKQYRKLA 15
 |||:|||||
 Db 398 dtrqgyrklik 408
 RESULT 14
 R27706
 ID R27706 standard; Protein; 591 AA.
 XX
 AC R27706;
 XX
 DT 16-MAR-1993 (first entry)
 DE
 DE Rat bone PTH/PTHrP receptor clone R15B prod.
 XX
 KW Parathyroid hormone; related protein; calcium; antagonist;
 KW antibodies; hypercalcaemia.
 XX
 OS Rattus rattus.
 XX
 PN W09217602-A.
 XX
 PD 15-OCT-1992.
 XX
 PF 06-APR-1992; 92WO-US02821.
 XX
 PR 05-APR-1991; 91US-0681702.
 PR 06-APR-1992; 92US-0864475.
 XX
 PA (GENO) GEN HOSPITAL CORP OFFICE TECHNOLOGY.
 XX
 XX Abou-samra A, Juppner H, Kronenberg HM, Potts JT, Schipani E;
 PI Segre GV;
 PI
 XX WPI; 1992-366271/44.
 DR N-ESDB; Q29606.
 DR
 XX New DNA encoding parathyroid hormone receptor. DNA and antibodies
 PT - for (differential) diagnosis of hypercalcaemia, and diagnosis
 PT and treatment of tumours
 XX
 PS Claim 20; Fig 3; 91pp; English.
 XX
 CC The rat bone parathyroid hormone/parathyroid hormone related
 CC protein (PTH/PTHrP) receptor protein sequence was deduced from
 CC clone R15B obtd. by screening a rat osteosarcoma (ROS) cell cDNA library

CC to isolate those expressing functionally intact PTH/PTHrP receptor
 CC proteins, performed according to Gearing et al., (EMBO J. 8: 3676,
 CC 1989), by identifying colonies capable of binding a suitable radio-
 CC labelled ligand. The protein may be used in a therapeutic compsn. to
 CC inhibit activation of PTH or PTHrP and thus reduce the level of calcium
 CC in the blood. Cpbs. capable of competing with PTH or PTHrP for binding
 CC can be identified using the protein and DNAs homologous to PTH DNA can
 CC be identified using fragments of the clone as probes. The sequence
 CC may be used for the prodn. of antibodies useful for the treatment,
 CC classification, prognosis and/or treatment of disorders related to
 CC the interaction between a cell receptor and a ligand such as in
 CC hypercalcaemia. See also R27704-16.
 XX
 SQ Sequence 591 AA;
 Query Match 56.4%; Score 44; DB 13; Length 591;
 Best Local Similarity 72.7%; Pred. No. 5.9;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 5 DTRKQYRKLA 15
 |||:|||||
 Db 398 dtrqgyrklik 408
 RESULT 15
 R92277
 ID R92277 standard; Protein; 591 AA.
 XX
 AC R92277;
 XX
 DT 18-MAY-1996 (first entry)
 DE
 DE Rat bone PTH/PTHrP receptor.
 XX
 KW Parathyroid hormone; receptor; parathormone; PTH;
 KW parathyroid hormone-related protein; PTHrP; calcium; homeostasis;
 KW hypercalcaemia; hypocalcaemia; cancer.
 XX
 OS Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT Region 1..192
 FT /label= Extracellular_region
 FT Region 193..211
 FT /label= Transmembrane_region
 FT Region 212..221
 FT /label= Intracellular_region
 FT Region 222..240
 FT /label= Transmembrane_region
 FT Region 241..299
 FT /label= Extracellular_region
 FT Region 300..316
 FT /label= Transmembrane_region
 FT Region 317..325
 FT /label= Intracellular_region
 FT Region 326..342
 FT /label= Transmembrane_region
 FT Region 343..364
 FT /label= Extracellular_region
 FT Region 365..383
 FT /label= Transmembrane_region
 FT Region 384..408
 FT /label= Intracellular_region
 FT Region 409..428
 FT /label= Transmembrane_region
 FT Region 429..444
 FT /label= Intracellular_region
 FT Region 445..463
 FT /label= Transmembrane_region
 FT Region 464..591
 FT /label= Intracellular_region
 XX


```

.DT 18-MAY-1996 (first entry)
XX
DE Human kidney PTH/PTHrP receptor.
XX
KW Parathyroid hormone; receptor; parathormone; PTH;
KW parathyroid hormone-related protein; PTHrP; calcium; homeostasis;
KW hypercalcaemia; hypocalcaemia; cancer.
XX
OS Homo sapiens.
XX
PN US5494806-A.
XX
PD 27-FEB-1996.
XX
PF 05-APR-1991; 91US-0681702.
XX
PR 06-APR-1992; 92US-0864475.
PR 05-APR-1991; 91US-0681702.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
PI Schipani E, Segre GV;
XX
DR WPI; 1996-139028/14.
DR N-PSDB; T15948.
XX
PT DNA encoding vertebrate parathyroid hormone receptor - useful for
PT diagnosis and treatment of e.g. hypercalcaemia, hypo-calcaemia,
PT cancer etc.
XX
PS Claim 1; Fig 6A-6G; 64pp; English.
XX
CC A human parathyroid hormone/parathyroid hormone-related protein
CC (PTH/PTHrP) receptor (R92278) is encoded by cDNA clone HK-1
CC (T15948) isolated from a human kidney cDNA library. The receptor
CC induces an increase in intracellular cAMP and intracellular free
CC calcium when challenged by PTH or PTHrP. Recombinant receptor can
CC be produced in vector/host cell systems and used in the treatment.
CC diagnosis or prognosis of e.g. hypercalcaemia and hypocalcaemia, to
CC screen for (ant)agonists and to raise antibodies. Host cells
CC expressing the receptor can be used for diagnostic measurement of
CC PTH serum levels.
XX
SQ Sequence 593 AA;

Query Match 60.3%; Score 47; DB 17; Length 593;
Best Local Similarity 81.8%; Pred. No. 1.8;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 DTRKQYRKAK 15
   |||:|||||
DB 398 dtrgqyrklk 408

RESULT 12
W73317
ID W73317 standard; Protein; 593 AA.
XX
AC W73317;
XX
DT 08-FEB-1999 (first entry)
DE Human Parathyroid hormone receptor.
XX
KW Parathyroid hormone receptor; PTH receptor; antibody; therapy;
KW PTH-related hypercalcaemia; human.
XX
OS Homo sapiens.
XX
PN US5840853-A.
XX
PT New DNA encoding parathyroid hormone receptor, DNA and antibodies

PD 24-NOV-1998.
XX
PF 06-JUN-1995; 95US-0471494.
XX
PR 06-APR-1992; 92US-0864475.
PR 05-APR-1991; 91US-0681702.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
PI Schipani E, Segre GV;
XX
DR WPI; 1999-034124/03.
DR N-PSDB; V08391.
XX
PT Antibody to parathyroid hormone receptor - for diagnostic or
PT therapeutic use
XX
PS Claim 7; Fig 6; 63pp; English.
XX
CC This sequence represents the human parathyroid hormone (PTH) receptor
CC which is targeted by the antibody of the invention. The antibody
CC of the invention is immunoreactive with naturally occurring human, rat or
CC opossum PTH receptor. The antibody is useful for treating disorders
CC characterised by overstimulation of PTH receptors by their ligand and for
CC the diagnosis of PTH-related hypercalcaemia.
XX
SQ Sequence 593 AA;

Query Match 60.3%; Score 47; DB 20; Length 593;
Best Local Similarity 81.8%; Pred. No. 1.8;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 DTRKQYRKAK 15
   |||:|||||
DB 398 dtrgqyrklk 408

RESULT 13
R27707
ID R27707 standard; Protein; 614 AA.
XX
AC R27707;
XX
DT 16-MAR-1993 (first entry)
DE Human kidney PTH/PTHrP receptor.
XX
KW Parathyroid hormone; related protein; calcium; antagonist;
KW antibodies; hypercalcaemia.
XX
OS Homo sapiens.
XX
PN WO9217602-A.
XX
PD 15-OCT-1992.
XX
PF 06-APR-1992; 92WO-US02821.
XX
PR 05-APR-1991; 91US-0681702.
PR 06-APR-1992; 92US-0864475.
XX
PA (GEHO ) GEN HOSPITAL CORP OFFICE TECHNOLOGY.
XX
PI Abou-samra A, Juppner H, Kronenberg HM, Potts JT, Schipani E;
PI Segre GV;
XX
DR WPI; 1992-366271/44.
DR N-PSDB; Q29607.
XX
PT New DNA encoding parathyroid hormone receptor, DNA and antibodies

```

CC Level of calcium in the blood. Cpd. capable of competing with PTH
 CC or PTHrP for binding can be identified using the protein prod. and
 CC DNAs homologous to PTH DNA can be identified using fragments of the
 CC clone as probes. The sequence may be used for the prodn. of antibodies
 CC useful for the treatment, classification, prognosis and/or treatment of
 CC disorders related to the interaction between a cell receptor and a
 CC ligand such as in hypercalcaemia. See also R27704-16.
 XX
 SQ Sequence 585 AA;

Query Match 60.3%; Score 47; DB 13; Length 585;
 Best Local Similarity 81.8%; Pred. No. 1.8;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 DTRKQYRKLA 15
 |||:|||||
 Db 392 dtrqgyrklik 402

RESULT 9
 R92276
 ID R92276 standard; Protein; 585 AA.
 AC R92276;
 DT 18-MAY-1996 (first entry)
 DE Opossum kidney PTH/PTHrP receptor.
 KW Parathyroid hormone; receptor; parathormone; PTH;
 KW parathyroid hormone-related protein; PTHrP; calcium; homeostasis;
 KW hypercalcaemia; hypocalcaemia; cancer; opossum.
 XX
 OS Didelphis virginiana.
 XX
 PN US5494806-A.
 XX
 PD 27-FEB-1996.
 XX
 PF 05-APR-1991; 91US-0681702.
 XX
 PR 06-APR-1992; 92US-0864475.
 PR 05-APR-1991; 91US-0681702.
 XX
 PA (GEO) GEN HOSPITAL CORP.
 XX
 PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;
 XX
 DR WPI; 1996-139028/14.
 DR N-PSDB; T15946.
 XX
 PT DNA encoding vertebrate parathyroid hormone receptor - useful for
 PT diagnosis and treatment of e.g. hypercalcaemia, hypo-calcaemia,
 PT cancer etc.
 PS
 XS Claim 1; Fig 2A-2E; 64pp; English.

XX Parathyroid hormone/parathyroid hormone-related protein (PTH/PTHrP)
 CC receptors (R92275 and R92276) are encoded by cDNA clones OK-H (T15945)
 CC and OK-O (T15946), respectively, isolated from an opossum kidney (OK)
 CC cell cDNA library. OK-H and OK-O may be the products of 2 separate
 CC genes or of a laboratory artifact. The receptor induces an increase
 CC in intracellular cAMP and calcium when challenged with PTH or PTHrP.
 CC Recombinant receptors can be produced in vector/host cell systems and
 CC used in the treatment, diagnosis or prognosis of e.g. hypercalcaemia
 CC and hypocalcaemia, to screen for (ant)agonists and to raise antibodies.
 CC Host cells expressing the receptor can be used for diagnostic
 CC measurement of PTH serum levels.
 XX
 SQ Sequence 585 AA;

Query Match 60.3%; Score 47; DB 17; Length 585;
 Best Local Similarity 81.8%; Pred. No. 1.8;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 DTRKQYRKLA 15
 |||:|||||
 Db 392 dtrqgyrklik 402

RESULT 10
 W73315
 ID W73315 standard; Protein; 585 AA.
 XX
 AC W73315;
 DT 08-FEB-1999 (first entry)
 DE Parathyroid hormone receptor OK-O.
 XX
 KW Parathyroid hormone receptor; PTH receptor; antibody; therapy;
 KW PTH-related hypercalcaemia; opossum.
 XX
 OS Didelphis virginiana.
 XX
 PN US5840853-A.
 XX
 PD 24-NOV-1998.
 XX
 PF 06-JUN-1995; 95US-0471494.
 XX
 PR 06-APR-1992; 92US-0864475.
 PR 05-APR-1991; 91US-0681702.
 PR 06-JUN-1995; 95US-0471494.
 XX
 PA (GEO) GEN HOSPITAL CORP.
 XX
 PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;
 XX
 DR WPI; 1999-034124/03.
 DR N-PSDB; V08389.
 XX
 PT Antibody to parathyroid hormone receptor - for diagnostic or
 PT therapeutic use
 XX
 PS Claim 6; Fig 2; 63pp; English.

XX This sequence represents the opossum parathyroid hormone (PTH) receptor
 CC OK-O, which is targeted by the antibody of the invention. The antibody
 CC of the invention is immunoreactive with naturally occurring human, rat or
 CC opossum PTH receptor. The antibody is useful for treating disorders
 CC characterised by overstimulation of PTH receptors by their ligand and for
 CC the diagnosis of PTH-related hypercalcaemia.
 XX
 SQ Sequence 585 AA;

Query Match 60.3%; Score 47; DB 20; Length 585;
 Best Local Similarity 81.8%; Pred. No. 1.8;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 DTRKQYRKLA 15
 |||:|||||
 Db 392 dtrqgyrklik 402

RESULT 11
 R92278
 ID R92278 standard; Protein; 593 AA.
 XX
 AC R92278;
 XX

```

PR 06-APR-1992; 92US-0864475.
PR 05-APR-1991; 91US-0681702.
XX
XX (GEO ) GEN HOSPITAL CORP.
XX
XX Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
XX Schipani E, Segre GV;
XX
XX WPI; 1996-139028/14.
XX N-PSDB; T15945.
XX
XX DNA encoding vertebrate parathyroid hormone receptor - useful for
XX diagnosis and treatment of e.g. hypercalcaemia, hypo-calcaemia,
XX cancer etc.
XX
XX Claim 1; Fig 1A-1E; 64pp; English.
XX
XX Parathyroid hormone/parathyroid hormone-related protein (PTH/PTHrP)
XX receptors (R92275 and R92776) are encoded by cDNA clones OK-H (T15945)
XX and OK-O (T15946), respectively, isolated from an opossum kidney (OK)
XX cell cDNA library. OK-H and OK-O may be the products of 2 separate
XX genes or of a laboratory artifact. The receptor induces an increase
XX in intracellular cAMP and calcium when challenged with PTH or PTHrP.
XX Recombinant receptors can be produced in vector/host cell systems and
XX used in the treatment, diagnosis or prognosis of e.g. hypercalcaemia
XX and hypocalcaemia, to screen for (ant)agonists and to raise antibodies.
XX Host cells expressing the receptor can be used for diagnostic
XX measurement of PTH serum levels.
XX
XX Sequence 515 AA;

Query Match 60.3%; Score 47; DB 17; Length 515;
Best Local Similarity 81.8%; Pred. No. 1.5;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRKQYRKLA 15
   |||:|||||
Db 392 dtrqyrrkllk 402

RESULT 7
W73314
ID W73314 standard; Protein; 515 AA.
XX
XX AC W73314;
XX
XX DT 08-FEB-1999 (first entry)
XX
XX DE Parathyroid hormone receptor OK-H.
XX
XX KW Parathyroid hormone receptor; PTH receptor; antibody; therapy;
XX PTH-related hypercalcaemia; opossum.
XX
XX OS Didelphis virginiana.
XX
XX PN US5840853-A.
XX
XX PD 24-NOV-1998.
XX
XX PF 06-JUN-1995; 95US-0471494.
XX
XX PR 06-APR-1992; 92US-0864475.
XX 05-APR-1991; 91US-0681702.
XX 06-JUN-1995; 95US-0471494.
XX
XX PA (GEO ) GEN HOSPITAL CORP.
XX
XX PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
XX Schipani E, Segre GV;
XX
XX WPI; 1999-034124/03.
XX N-PSDB; V08388.

Query Match 60.3%; Score 47; DB 20; Length 515;
Best Local Similarity 81.8%; Pred. No. 1.5;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DTRKQYRKLA 15
   |||:|||||
Db 392 dtrqyrrkllk 402

RESULT 8
R27705
ID R27705 standard; Protein; 585 AA.
XX
XX AC R27705;
XX
XX DT 16-MAR-1993 (first entry)
XX
XX DE Opossum kidney PTH/PTHrP receptor prod. from clone OK-O.
XX
XX KW Parathyroid hormone; related protein; calcium; antagonist;
XX antibodies; hypercalcaemia.
XX
XX OS Didelphis virginiana.
XX
XX PN W09217602-A.
XX
XX PD 15-OCT-1992.
XX
XX PF 06-APR-1992; 92WO-US02821.
XX
XX PR 05-APR-1991; 91US-0681702.
XX 06-APR-1992; 92US-0864475.
XX
XX PA (GEO ) GEN HOSPITAL CORP OFFICE TECHNOLOGY.
XX
XX PI Abou-samra A, Juppner H, Kronenberg HM, Potts JT, Schipani E;
XX Segre GV;
XX
XX WPI; 1992-366271/44.
XX N-PSDB; Q29605.
XX
XX PT New DNA encoding parathyroid hormone receptor, DNA and antibodies
XX - for (differential) diagnosis of hypercalcaemia, and diagnosis
XX and treatment of tumours
XX
XX PS Disclosure; Fig 2; 91pp; English.
XX
XX The parathyroid hormone/parathyroid hormone related protein (PTH/PTHrP)
XX receptor protein sequence was deduced from the DNA sequence of the
XX clone OK-O, isolated from opossum kidney (OK) cells. The clone OK-O
XX is identical to the OK-H clone except at the C-terminal tail as OK-O
XX encodes a 585 amino acid protein, OK-H encodes one of 515 amino acids.
XX The difference is attributed to a single nucleotide deleted in the OK-H
XX sequence causing a frame shift and an earlier stop codon. It is not
XX known whether OK-O and OK-H represent prods. of two separate genes or
XX are a laboratory artifact. The protein may be used in a therapeutic
XX compsn. to inhibit activation of PTH or PTHrP and thus reduce the

```

QY 5 DTRKQYRKLA 15
 III:IIII I
 Db 15 dtrgqyrkllk 25

RESULT 4
 W73326
 ID W73326 standard; peptide; 25 AA.

XX AC W73326;
 XX AC
 DT 08-FEB-1999 (first entry)
 XX 08-FEB-1999 (first entry)

DE Parathyroid hormone receptor fragment RPI-9.

XX Parathyroid hormone receptor; PTH receptor; antibody; therapy;
 KW PTH-related hypercalcaemia; opossum; rat; human.

XX Synthetic.

XX US5840853-A.

PN 24-NOV-1998.

XX 06-JUN-1995; 95US-0471494.

XX 06-APR-1992; 92US-0864475.

PR 05-APR-1991; 91US-0681702.

PR 06-JUN-1995; 95US-0471494.

XX (GEO) GEN HOSPITAL CORP.

PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;

PI Schipani E, Segre GV;

XX WPI; 1999-034124/03.

XX Antibody to parathyroid hormone receptor - for diagnostic or

PT therapeutic use

XX Disclosure; Column 19; 63pp; English.

XX This sequence is a fragment of a opossum parathyroid hormone (PTH)
 CC receptor which is targeted by the antibody of the invention. The antibody
 CC of the invention is immunoreactive with naturally occurring human, rat or
 CC opossum PTH receptor. The antibody is useful for treating disorders
 CC characterised by overstimulation of PTH receptors by their ligand and for
 CC the diagnosis of PTH-related hypercalcaemia.

XX Sequence 25 AA;

Query Match 60.3%; Score 47; DB 20; Length 25;
 Best Local Similarity 81.8%; Pred. No. 0.073;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 DTRKQYRKLA 15
 III:IIII I
 Db 15 dtrgqyrkllk 25

RESULT 5
 R27704
 ID R27704 standard; Protein; 515 AA.

XX AC R27704;

XX 16-MAR-1993 (first entry)

DE Opossum kidney PTH/PTHrP receptor prep. from clone OK-H.

XX Parathyroid hormone; related protein; calcium; antagonist;
 KW antibodies; hypercalcaemia.

XX Didelphis virginiana.
 OS WO9217602-A.
 XX PD 15-OCT-1992.

XX 06-APR-1992; 92WO-US02821.

XX 05-APR-1991; 91US-0681702.

PR 06-APR-1992; 92US-0864475.

XX (GEO) GEN HOSPITAL CORP OFFICE TECHNOLOGY.

XX Abou-samra A, Juppner H, Kronenberg HM, Potts JT, Schipani E;

PI Segre GV;

XX WPI; 1992-366271/44.

XX N-PSDB; Q29604.

XX New DNA encoding parathyroid hormone receptor, DNA and antibodies

PT - for (differential) diagnosis of hypercalcaemia, and diagnosis

PT and treatment of tumours

XX Disclosure; Fig 1; 91pp; English.

XX The parathyroid hormone/parathyroid hormone related protein (PTH/PTHrP)

CC receptor protein sequence was deduced from the DNA sequence of the

CC clone OK-H, isolated from opossum kidney (OK) cells. The protein

CC may be used in a therapeutic compsn. to inhibit activation of PTH or

CC PTHrP and thus reduce the level of calcium in the blood. Cpd.

CC capable of competing with PTH or PTHrP for binding can be identified

CC using the protein prod. and DNAs homologous to PTH DNA can be

CC identified using fragments of the clone as probes. The protein

CC may be used for the prodn. of antibodies useful for the treatment,

CC classification, prognosis and/or treatment of disorders related to

CC the interaction between a cell receptor and a ligand such as in

CC hypercalcaemia. See also R27705-16.

XX Sequence 515 AA;

Query Match 60.3%; Score 47; DB 13; Length 515;
 Best Local Similarity 81.8%; Pred. No. 1.5;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 DTRKQYRKLA 15
 III:IIII I

Db 392 dtrgqyrkllk 402

RESULT 6

R92275

ID R92275 standard; Protein; 515 AA.

XX AC R92275;

XX 18-MAY-1996 (first entry)

XX Opossum kidney PTH/PTHrP receptor.

XX Parathyroid hormone; receptor; parathormone; PTH;

KW parathyroid hormone-related protein; PTHrP; calcium; homeostasis;

KW hypercalcaemia; hypocalcaemia; cancer; opossum.

XX Didelphis virginiana.

XX US5494806-A.

XX 27-FEB-1996.

XX 05-APR-1991; 91US-0681702.

XX

PT hyper-calcaemia, hypo- or hyper-phosphatemia, kidney stones, etc
 PS Claim 9; Fig 1A-E; 62pp; English.
 XX
 CC A novel 7-transmembrane receptor (WL2695) has been identified as a
 CC human G-protein parathyroid hormone (PTH) receptor, designated
 CC HLTG74. It shows 48.2% homology to the human PTH receptor. Its
 CC amino acid sequence was deduced from a cDNA clone (T59619) isolated
 CC from a human T cell lymphoma tissue cDNA library. Recombinant
 CC HLTG74 can be produced in transfected host cells and used to
 CC screen for (antagonist) cpds. Agonists may be used to prevent or
 CC treat e.g. hypocalcaemia, hyperphosphataemia, hypoparathyroidism
 CC and chronic tetany by stimulating an increase in serum calcium
 CC levels. Antagonists can be used to inhibit the receptor e.g. for
 CC the treatment of osteoporosis, hypercalcaemia, hypoparathyroidism,
 CC hypophosphataemia, kidney stone, nephrolisis.
 XX
 SQ Sequence 541 AA;

Query Match 100.0%; Score 78; DB 18; Length 541;
 Best Local Similarity 100.0%; Pred. No. 6e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVGHDKQYRKLA 15
 |||||
 DB 349 avghdtrkqrklak 363

RESULT 2
 R27716
 ID R27716 standard; peptide; 25 AA.
 AC R27716;
 XX
 DT 16-MAR-1993 (first entry)
 DE PTH/PTHrP receptor fragment.
 KW Parathyroid hormone; related protein; calcium; antagonist;
 KW antibodies; hypercalcaemia; intracellular domain.
 XX
 OS Synthetic.
 XX WO9217602-A.
 XX 15-OCT-1992.
 XX
 PF 06-APR-1992; 92WO-US02821.
 XX
 PR 05-APR-1991; 91US-0681702.
 PR 06-APR-1992; 92US-0864475.
 XX
 PA (GEHO) GEN HOSPITAL CORP OFFICE TECHNOLOGY.
 XX
 PI Abou-samra A, Juppner H, Kronenberg HM, Potts JT, Schipani E;
 PI Segre GV;
 XX WPI; 1992-366271/44.
 DR
 XX New DNA encoding parathyroid hormone receptor, DNA and antibodies
 PT - for (differential) diagnosis of hypercalcaemia, and diagnosis
 PT and treatment of tumours
 XX
 PS Claim 25; Page 5; 91pp; English.
 XX

The peptide sequence shown represents an intracellular fragment of
 CC parathyroid hormone/parathyroid hormone related protein
 CC (PTH/PTHrP) receptor protein. The peptide is capable of binding
 CC PTH or PTHrP and acting as an antagonist of these cpds. The
 CC peptide may be used to inhibit activation of PTH or PTHrP and thus
 CC reduce the level of calcium in the blood. Cpds. capable of competing
 CC with PTH or PTHrP for binding can be identified using the protein prod.

CC and DNAs homologous to PTH DNA can be identified using fragments of the
 CC clone as probes. The sequence may be used for the prodn. of antibodies
 CC useful for the treatment, classification, prognosis and/or treatment of
 CC disorders related to the interaction between a cell receptor and a
 CC ligand such as in hypercalcaemia. See also R27704-15.
 XX
 SQ Sequence 25 AA;

Query Match 60.3%; Score 47; DB 13; Length 25;
 Best Local Similarity 81.8%; Pred. No. 0.073;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 DTRKQYRKLA 15
 |||||
 DB 15 dtrqqrklk 25

RESULT 3
 R92287
 ID R92287 standard; Peptide; 25 AA.
 XX
 AC R92287;
 XX
 DT 18-MAY-1996 (first entry)
 XX PTH/PTHrP intracellular region fragment RPI-9.
 DE
 KW Parathyroid hormone; receptor; parathormone; PTH;
 KW parathyroid hormone-related protein; PTHrP; calcium; homeostasis;
 KW hypercalcaemia; hypocalcaemia; cancer.
 XX
 OS Synthetic.
 XX US5494806-A.
 XX 27-FEB-1996.
 XX
 PF 05-APR-1991; 91US-0681702.
 XX
 PR 06-APR-1992; 92US-0864475.
 PR 05-APR-1991; 91US-0681702.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;
 XX WPI; 1996-139028/14.
 DR
 XX DNA encoding vertebrate parathyroid hormone receptor - useful for
 PT diagnosis and treatment of e.g. hypercalcaemia, hypo-calcaemia,
 PT cancer etc.
 XX
 PS Disclosure; Column 3; 64pp; English.
 XX
 CC Parathyroid hormone (PTH) receptor fragments, including those
 CC (R92279-84) based on the PTH extracellular region and those
 CC (R92285-87) based on the intracellular domain, are produced by
 CC incorporating encoding DNA sequences into a vector, and
 CC culturing cells transformed by the vector. The peptides can be
 CC used to raise antibodies. The peptides and antibodies are useful
 CC in the treatment, diagnosis or prognosis of e.g. hypercalcaemia and
 CC hypocalcaemia, and can also be used to screen for (ant)agonists
 CC of therapeutic appln.
 XX
 SQ Sequence 25 AA;

Query Match 60.3%; Score 47; DB 17; Length 25;
 Best Local Similarity 81.8%; Pred. No. 0.073;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:51:44 ; Search time 138.73 Seconds
(without alignments)
3.697 Million cell updates/sec

Title: US-09-236-468A-2_COPY_349_363
Perfect score: 78
Sequence: 1 AVGHDRKQYKLA 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_36.*

1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	100.0	541	18	W12695
2	47	60.3	25	13	R27716
3	47	60.3	25	17	R92287
4	47	60.3	25	20	W73326
5	47	60.3	515	13	R27704
6	47	60.3	515	17	R92275
7	47	60.3	515	20	W73314
8	47	60.3	585	13	R27705
9	47	60.3	585	17	R92276
10	47	60.3	585	20	W73315
11	47	60.3	593	17	R92278
12	47	60.3	593	20	W73317

13	47	60.3	614	13	R27707
14	44	56.4	591	13	R27706
15	44	56.4	591	17	R92277
16	44	56.4	591	20	W73316
17	42	53.8	543	20	W70558
18	41	52.6	257	17	W02162
19	41	52.6	269	17	R98459
20	41	52.6	269	17	W02161
21	40	51.3	223	17	R90880
22	39	50.0	1432	18	Y14517
23	38	48.7	415	20	Y07061
24	38	48.7	431	19	W80310
25	38	48.7	431	20	W92974
26	38	48.7	437	16	R70136
27	38	48.7	437	19	W80309
28	38	48.7	437	20	W92973
29	38	48.7	438	19	W80308
30	38	48.7	438	20	W92972
31	38	48.7	785	21	Y81536
32	37	47.4	75	21	Y55671
33	37	47.4	131	20	Y12477
34	37	47.4	371	20	W90181
35	37	47.4	417	21	Y55646
36	37	47.4	631	12	R13509
37	37	47.4	1417	18	W31548
38	37	47.4	1417	18	W31550
39	37	47.4	1417	18	W31551
40	37	47.4	1418	18	W15264
41	36	46.2	104	18	Y11178
42	36	46.2	254	19	Y85806
43	36	46.2	286	19	W38535
44	36	46.2	317	20	Y74126
45	36	46.2	330	20	W94066

ALIGNMENTS

RESULT 1

W12695
ID W12695 standard; Protein; 541 AA.

XX W12695;

XX AC
DT 31-MAY-1997 (first entry)

XX G-protein parathyroid hormone receptor HLTGDG74.

DE G-protein parathyroid hormone receptor; HLTGDG74; parathormone; PTH;
KW calcium; signal transduction; agonist; antagonist; hypocalcaemia;
KW hyperphosphataemia; hypoparathyroidism; chronic tetany;
KW osteoporosis; hypercalcaemia; hypoparathyroidism; hypophosphataemia;
KW kidney stone; nephrolithiasis; therapy; diagnosis.

XX Homo sapiens.

XX WO9639433-A1.

XX PD 12-DEC-1996.

XX PF 05-JUN-1995; 95WO-US07085.

XX PR 05-JUN-1995; 95WO-US07085.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Li Y, Rosen CA, Ruben SM, Soppet DR;

XX WPI: 1997-043068/04.

XX N-PSDB; T59619.

XX Human G-protein parathyroid hormone receptor, HLTGDG74 - used to
PT identify (ant)agonists, used in the treatment of hypo- or

Human kidney PTH/P
Rat bone PTH/PTHrP
Rat bone PTH/PTHrP
Parathyroid hormon
Renal Cancer assoc
Lactococcus lactis
m6A-methyltransfer
Lactococcus lactis
Mouse cysteine str
Human WRN gene pro
Renal cancer assoc
Human PACAP/VIP R-
Human PACAP/VIP R2
Rat vasoactive int
Rat PACAP/VIP R-2
Rat PACAP/VIP R-2
Human PACAP/VIP R-
Human PACAP/VIP R-
Streptococcus pneu
S. cerevisiae ycfB
Human 5' EST seque
Pantoea citrea cry
S. cerevisiae ycfB
P. denitrificans CO
Bloom's syndrome B
Bloom's syndrome B
Bloom's syndrome B
Bloom syndrome act
S. pneumoniae 30S
S. pneumoniae deri
S. pneumoniae 30S
Human prostate tum
Human DnaJ-like pr

Qy 1 SAGDIK 6
:|||||
Db 269 TAGDIK 274

Search completed: November 8, 2000, 08:56:10
Job time: 521 sec

DR PFAM; PF00106; adh_short; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 DR PRINTS; PR00081; GDRDH.
 DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
 SQ SEQUENCE 263 AA; 27814 MW; 07A9200A2653EB20 CRC64;

Query Match 89.7%; Score 26; DB 2; Length 263;
 Best Local Similarity 83.3%; Pred. No. 2.6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6
 :|||||
 Db 191 SAGDIR 196

RESULT 13

ID O93320 PRELIMINARY; PRT; 351 AA.
 AC O93320;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE P80 KATANIN (FRAGMENT).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-OVARY;
 RA McNally F.J., Thomas S.;
 RT "Katanin is Responsible for the M-Phase Microtubule-Severing Activity
 in Xenopus Eggs."; 0-0-0(1998).
 RL Mol. Biol. Cell 0:0-0(1998).
 DR EMBL; AF056021; AAC25113.1; -.
 FT NON_TER 1
 SQ SEQUENCE 351 AA; 38344 MW; 2BD4163F03787BA0 CRC64;

Query Match 89.7%; Score 26; DB 13; Length 351;
 Best Local Similarity 83.3%; Pred. No. 3.6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6
 :|||||
 Db 213 SSGDIK 218

RESULT 14

ID O29766 PRELIMINARY; PRT; 378 AA.
 AC O29766;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JUN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE GTP CYCLOHYDROLASE II (RIBA-1).
 GN AF0484.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 OC Archaeoglobus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE; 98049343.
 RA Klenk H.-P., Clayton R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Ketchum K.A., Dodson R.J., Graham D.E., Kyripides N.C.,
 RA Richardson D.L., Kerlavage A.R., Lee N.H., Sutton G.G., Gill S.,
 RA Fleischmann R.D., Quackenbush J., Adams M.D., Loftus B.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Lofus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,

RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaeon Archaeoglobus fulgidus."; 390:364-370(1997).
 RL Nature 390:364-370(1997).
 DR EMBL; AF001070; AAB90751.1; -.
 DR TIGR; AF0484; -.
 DR INTERPRO; IPR000422; -.
 DR INTERPRO; IPR000926; -.
 DR PFAM; PF00925; GTP_cyclohydro2; 1.
 DR PFAM; PF00926; DHBP_synthase; 1.
 DR PRODOM; PD003034; -. 1.
 KW Hypothetical protein; Hydrolase.
 SQ SEQUENCE 378 AA; 42271 MW; 5F63B95448817356 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 378;
 Best Local Similarity 83.3%; Pred. No. 3.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6
 :|||||
 Db 96 NAGDIK 101

RESULT 15

O77027 PRELIMINARY; PRT; 403 AA.
 ID O77027
 AC O77027;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE CUBITUS INTERRUPTUS PROTEIN (CI PROTEIN) (CUBITUS INTERRUPTUS DOMINANT
 DE PROTEIN) (FRAGMENT).
 DE CI OR CID.
 OS Drosophila yakuba (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 98278813.
 RA Takano T.S.;
 RT "Rate variation of DNA sequence evolution in the Drosophila
 lineages."; 149:959-970(1998).
 RL Genetics 149:959-970(1998).
 CC -!- FUNCTION: INVOLVED IN SEGMENT POLARITY. IS REQUIRED FOR THE NORMAL
 DEVELOPMENT OF THE POSTERIOR HALF OF EACH EMBRYONIC SEGMENT. EN
 CC PROTEIN DIRECTLY REPRESSSES CI EXPRESSION IN POSTERIOR COMPARTMENT
 CC CELLS (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -!- SIMILARITY: TO THE GLI-RELATED GROUP OF C2H2-TYPE ZINC-FINGERS
 CC PROTEINS.
 DR EMBL; AB005797; BAA33208.1; -.
 DR FLYBASE; FBgn022826; Dyak\ci.
 DR INTERPRO; IPR000822; -.
 DR PFAM; PF00096; zf-C2H2; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2; UNKNOWN_2.
 DR Developmental protein; Segmentation polarity protein; Zinc-finger;
 KW Metal-binding; DNA-binding; Repeat; Nuclear protein.
 FT NON_TER 1
 FT DOMAIN <1 83 ZINC-FINGERS (POTENTIAL).
 FT ZN_FING <1 21 C2H2-TYPE (POTENTIAL).
 FT ZN_FING 27 52 C2H2-TYPE (POTENTIAL).
 FT ZN_FING 58 83 C2H2-TYPE (POTENTIAL).
 FT NON_TER 403 403
 SQ SEQUENCE 403 AA; 44089 MW; 467706D7D7E64359 CRC64;

Query Match 89.7%; Score 26; DB 5; Length 403;
 Best Local Similarity 83.3%; Pred. No. 4.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DT	01-JUN-2000 (TrEMBLrel. 14, Last annotation update)	
DE	PUTATIVE TRANSCRIPTIONAL REGULATOR.	
GN	SCF55.06.	
OS	Streptomyces coelicolor.	
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;	
OC	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.	
RN	[1]	
RN	SEQUENCE FROM N.A.	
RC	STRAIN=A3(2);	
RA	Seeger K.J., Harris D.;	
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.	
RN	[2]	
RN	SEQUENCE FROM N.A.	
RC	STRAIN=A3(2);	
RA	Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;	
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.	
RN	[3]	
RN	SEQUENCE FROM N.A.	
RC	STRAIN=A3(2);	
RA	Redenbach M., Kieser H.M., Hopwood D.A.;	
RA	Kinashi H., Hopwood D.A.;	
RT	"A set of ordered cosmids and a detailed genetic and physical map for	
RT	the 8 Mb Streptomyces coelicolor A3(2) chromosome.";	
DR	Mol. Microbiol. 21:77-96(1996).	
DR	EMBL; AL13291; CAB61276.1; -.	
DR	INTERPRO; IPR000524; -.	
DR	PFAM; PF00392; gntK; 1.	
DR	PRINTS; PR00035; HTHGNTR.	
SO	SEQUENCE 253 AA; 27502 MW; 72A4CFDB26E561C9 CRC64;	
Query Match	89.7%;	Score 26; DB 2; Length 253;
Best Local Similarity	83.3%;	Pred. No. 2.5e+02;
Matches	5; Conservative	1; Mismatches 0; Indels 0; Gaps
Qy	1 SAGDIK 6	
Db		
	43 SAGDIR 48	
RESULT	12	
Q9RK87		
ID	Q9RK87 PRELIMINARY; PRT; 263 AA.	
AC	Q9RK87;	
DT	01-MAY-2000 (TrEMBLrel. 13, Created)	
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)	
DT	01-JUN-2000 (TrEMBLrel. 14, Last annotation update)	
DE	PUTATIVE OXIDOREDUCTASE.	
GN	SCF11.02.	
OS	Streptomyces coelicolor.	
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;	
OC	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.	
RN	[1]	
RN	SEQUENCE FROM N.A.	
RC	STRAIN=A3(2);	
RA	Seeger K.J., Harris D.;	
RL	Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.	
RN	[2]	
RN	SEQUENCE FROM N.A.	
RC	STRAIN=A3(2);	
RA	Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;	
RL	Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.	
RN	[3]	
RN	SEQUENCE FROM N.A.	
RC	STRAIN=A3(2);	
RA	Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,	
RA	Kinashi H., Hopwood D.A.;	
RT	"A set of ordered cosmids and a detailed genetic and physical map for	
RT	the 8 Mb Streptomyces coelicolor A3(2) chromosome.";	
DR	Mol. Microbiol. 21:77-96(1996).	
DR	EMBL; AL132662; CAB59579.1; -.	
DR	INTERPRO; IPR002198; -.	
DR	INTERPRO; IPR002347; -.	

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003421; AAF45636.1; -;
DR FLYBASE: FBgn0026871; EG:80H7.11.
SQ SEQUENCE 325 AA; 36997 MW; FAEA49BD69D18152 CRC64;

Query Match 93.1%; Score 27; DB 5; Length 325;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6
|||||
Db 73 SAGDLK 78

RESULT 6
Q9XZT3 PRELIMINARY; PRT; 352 AA.
ID Q9XZT3
AC Q9XZT3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE EG:80H7.11 PROTEIN.
GN EG:80H7.11.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidea; Drosophilidae; *Drosophila*.
RN [1]
RP SEQUENCE FROM N.A.
RA Mottier S., Cadieu E., Dreano S., Lelaure V., Galibert F.;
RT "Sequencing the distal X chromosome of *Drosophila melanogaster*.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Benos P.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL031027; CAB41537.1; -;
SQ SEQUENCE 352 AA; 40121 MW; A5DD5A7B829807B1 CRC64;

Query Match 93.1%; Score 27; DB 5; Length 352;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6
|||||
Db 99 SAGDLK 104

RESULT 7
Q9R9R6 PRELIMINARY; PRT; 464 AA.
ID Q9R9R6
AC Q9R9R6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE FLAH.
GN FLAH.
OS *Aeromonas caviae* (*Aeromonas* formicans).
OC Bacteria; Proteobacteria; gamma subdivision; *Aeromonas* group;
OC *Aeromonas*.
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=SCH3N;
RA Rabaan A.A., Shaw J.G.;
RT "Molecular Characterisation of the polar flagella of *Aeromonas*
RT *caviae*.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AFI98617; AAF1982.1; -;
SQ SEQUENCE 464 AA; 48814 MW; 91CD372F4879720C CRC64;

Query Match 93.1%; Score 27; DB 2; Length 464;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6
|||||
Db 255 SAGDLK 260

RESULT 8
O73636 PRELIMINARY; PRT; 868 AA.
ID O73636
AC O73636;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE PHEROMONE RECEPTOR.
GN CA02.1.
OS Fugu rubripes (Japanese pufferfish) (*Takifugu rubripes*).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Acanthopterygii;
OC Percomorpha; Tetraodontiformes; Tetraodontidae;
OC *Takifugu*.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98226788.
RA Naito T., Saito Y., Yamamoto J., Nozaki Y., Tomura K., Hazama M.,
RA Nakanishi S., Brenner S.;
RT "Putative pheromone receptors related to the Ca2+-sensing receptor in
RT *Fugu*.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:5178-5181(1998).
DR EMBL: AB008858; BAA26123.1; -;
DR INTERPRO: IPR000337; -;
DR INTERPRO: IPR001828; -;
DR INTERPRO: IPR002052; -;
DR PFAM: PF00003; 7tm_3; 1.
DR PFAM: PF01094; ANF_receptor; 1.
DR PRINTS: PR00248; GPCRMR.
DR PROSITE: PS00092; N6_MTASE; UNKNOWN_1.
DR PROSITE: PS00981; G_PROTEIN_RECEP_F3_3; 1.
KW Pheromone.
SQ SEQUENCE 868 AA; 95531 MW; 3FC66EB1E9972E01 CRC64;

Query Match 93.1%; Score 27; DB 13; Length 868;
Best Local Similarity 83.3%; Pred. No. 5.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6
|||||
Db 18 SAGDLK 23

RESULT 9
Q66063 PRELIMINARY; PRT; 95 AA.
ID Q66063
AC Q66063;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE URACYL-DNA GLYCOSYLASE (FRAGMENT).
OS Canine herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae.

```

DE BETA-KETOACYL-ACP SYNTHASE II.
GN CWRASIII.
OS Cuphea wrightii.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Rosidae; eurosids II; Myrtales;
OC Lythraceae; Cuphea.
RN [1]
RP SEQUENCE FROM N.A.
RA Slabaugh M.B., Leonard J.L., Knapp S.J.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U67317; AAB37271.1; -.
DR HSSP; P39435; IB3N.
DR MENDEL; 11694; Cupw; 1448; 11694.
DR INTERPRO; IPR000495; -.
DR INTERPRO; IPR000794; -.
DR PFAM; PF00109; ketoacyl-synt; 1.
DR PROSITE; PS00290; IG.MHC; UNKNOWN_1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
SQ SEQUENCE 540 AA: 57884 MW: C61B032B7BC114C5 CRC64;

Query Match 100.0%; Score 29; DB 10; Length 540;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
Db 435 SAGDIK 440

RESULT 3
O04430 PRELIMINARY; PRT; 133 AA.
AC O04430;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE EXTRACELLULAR INSOLUBLE CYSTATIN.
GN EIC.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Asteridae; euasterids II; Apiales;
OC Apiaceae; Daucus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=US-HARUMAKIGOSUN;
RX MEDLINE; 97320465.
RA Ojima A., Shiota H., Higashi K., Kamada H., Shimma Y., Wada M.,
RA Satoh S.;
RT "An extracellular insoluble inhibitor of cysteine proteinases in cell
RT cultures and seeds of carrot."
RL Plant Mol. Biol. 34:99-109(1997).
DR EMBL; D85623; BAA20464.1; -.
DR MENDEL; 16294; Dauca; 120; 16294.
DR INTERPRO; IPR000010; -.
DR PROSITE; PS00287; CYSTATIN; UNKNOWN_1.
SQ SEQUENCE 133 AA: 14091 MW: 0587253521B5EF07 CRC64;

Query Match 93.1%; Score 27; DB 10; Length 133;
Best Local Similarity 83.3%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
Db 73 SAGDIK 78

RESULT 4
O96380 PRELIMINARY; PRT; 185 AA.
ID O96380;
AC O96380;
DT 01-MAY-1999 (TREMBLrel. 10, Created)

```

```

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE THIOREDOXIN PEROXIDASE (FRAGMENT).
OS Echinococcus granulosus.
OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
OC Cyclophyllidae; Taeniidae; Echinococcus.
RN [1]
RP SEQUENCE FROM N.A.
RA Salinas G., Fernandez V., Fernandez C., Selkirk M.E.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF034959; AAO2002.1; -.
DR HSSP; P30041; IPRX.
DR INTERPRO; IPR000866; -.
DR PFAM; PF00578; AhpC-TSA; 1.
DR Peroxidase.
KW NON_TER.
FT NON_TER.
SQ SEQUENCE 185 AA: 20635 MW: D531ACD25DA811F9 CRC64;

Query Match 93.1%; Score 27; DB 5; Length 185;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
Db 175 SAGDIK 180

RESULT 5
Q9W584 PRELIMINARY; PRT; 325 AA.
ID Q9W584;
AC Q9W584;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE EG:80H7.11 PROTEIN.
GN EG:80H7.11.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE; 20196006.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:56:07 ; Search time 152.43 seconds
(without alignments)
3.675 Million cell updates/sec

Title: US-09-236-468A-2_COPY_310_315

Perfect score: 29

Sequence: 1 SAGDIK 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 297973

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL14:*

- 1: sp-archaea:*
- 2: sp-bacteria:*
- 3: sp-fungi:*
- 4: sp-human:*
- 5: sp-invertebrate:*
- 6: sp-mammal:*
- 7: sp-mhc:*
- 8: sp-organelle:*
- 9: sp-phage:*
- 10: sp-plant:*
- 11: sp-rodent:*
- 12: sp-virus:*
- 13: sp-vertebrate:*
- 14: sp-unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	29	100.0	307	10 Q9SZF0	Q9szf0 arabidopsis
2	29	100.0	540	10 P93118	P93118 cuphea wrig
3	27	93.1	133	10 O04430	O04430 daucus caro
4	27	93.1	185	5 O96380	O96380 echinococcu
5	27	93.1	325	5 Q9W584	Q9W584 drosophila
6	27	93.1	352	5 Q9XZ73	Q9xzt3 drosophila
7	27	93.1	464	2 Q9R9R6	Q9r9r6 aeromonas c
8	27	93.1	868	13 Q73636	Q73636 fugu rubrip
9	26	89.7	95	12 Q66063	Q66063 canine herp
10	26	89.7	253	2 Q9RX88	Q9rx88 deinoococcus
11	26	89.7	253	2 Q9RJ08	Q9rjq8 streptomyce
12	26	89.7	263	2 Q9RK87	Q9rk87 streptomyce
13	26	89.7	351	13 Q93320	Q93320 xenopus lae
14	26	89.7	378	1 Q29766	Q29766 archaeglob
15	26	89.7	403	5 Q77027	Q77027 drosophila
16	26	89.7	744	5 Q9U016	Q9u0i6 plasmodium
17	26	89.7	766	13 P79994	P79994 gallus galli
18	26	89.7	949	5 Q9W430	Q9w4j0 drosophila
19	26	89.7	958	5 Q77274	Q77274 drosophila

20	26	89.7	958	5 Q9W007	Q9w0q7 drosophila
21	26	89.7	1168	2 P74359	P74359 synecocyst
22	26	89.7	2831	2 O85166	O85166 bruceella ab
23	26	89.7	7829	5 Q18559	Q18559 caenorhabdi
24	25	86.2	102	5 O44471	O44471 caenorhabdi
25	25	86.2	109	2 Q9ZC08	Q9zc08 streptomyc
26	25	86.2	114	2 Q9RTA9	Q9rtja9 streptomyc
27	25	86.2	116	2 O66151	O66151 actinobacill
28	25	86.2	127	2 Q9RLC9	Q9rlc9 pseudomonas
29	25	86.2	135	6 Q28968	Q28968 sus scrofa
30	25	86.2	136	6 Q95258	Q95258 sus scrofa
31	25	86.2	142	6 Q9XS14	Q9xs14 sus scrofa
32	25	86.2	154	6 O46381	O46381 oryctolagus
33	25	86.2	156	8 O20623	O20623 physarum po
34	25	86.2	180	2 Q51889	Q51889 proteus mir
35	25	86.2	180	2 Q51931	Q51931 proteus mir
36	25	86.2	190	12 Q9W8B1	Q9w8b1 human immun
37	25	86.2	193	6 Q28530	Q28530 macropus to
38	25	86.2	213	10 O49929	O49929 pisum sativ
39	25	86.2	218	11 O64401	O64401 cricetus
40	25	86.2	218	11 Q60466	Q60466 cricetus
41	25	86.2	218	13 Q9W719	Q9w719 gallus gall
42	25	86.2	226	2 Q59559	Q59559 mycoplasma
43	25	86.2	240	5 Q21559	Q21559 caenorhabdi
44	25	86.2	241	2 O86451	O86451 pseudomonas
45	25	86.2	258	10 Q39626	Q39626 cucumis sat

ALIGNMENTS

RESULT 1

Q9SZF0 ID Q9SZF0 PRELIMINARY; PRT; 307 AA.
AC Q9SZF0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE FORMAMIDASE-LIKE PROTEIN.
GN F19F18.50.

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Rosidae; eurosids II; Brassicales;
OC Brassicaceae; Arabidopsis.

RN [1]
RP SEQUENCE FROM N.A.

RA Bevan M., Koetter P., Hempel S., Entian K.-D., Bancroft I.,

RA Mewes H.W., Mayer K.F.X., Schueller C.;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

RN [2]
RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL035605; CAB38295.1; -

SQ SEQUENCE 307 AA; 33917 MW; 50F87F2F741228F4 CRC64;

Query Match Best Local Similarity 100.0%; Score 29; DB 10; Length 307;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6

Db 62 SAGDIK 67

RESULT 2

P93118 ID P93118 PRELIMINARY; PRT; 540 AA.
AC P93118;
DT 01-MAY-1997 (TREMBLrel. 03, Created)

DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)

DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

Best Local Similarity 83.3%; Pred. No. 4.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
|||||
Db 159 SAGDIR 164

RESULT 12
S76195
hypothetical protein sll1527 - *Synechocystis* sp. (strain PCC 6803)
C:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S76195
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
s.
A:Reference number: S74322; MUID:97061201
A:Accession: S76195
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1168 <KAN>
A:Cross-references: EMBL:D90914; GB:AB001339; NID:g1653477; PIDN:BAAL8454.1; PID:g165354
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: *Synechocystis* hypothetical protein sll1527

Query Match 89.7%; Score 26; DB 2; Length 1168;
Best Local Similarity 83.3%; Pred. No. 6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
|||||
Db 49 SAGDIK 54

RESULT 13
T31419
cyclic beta 1-2 glucan synthetase - *Brucella abortus*
C:Species: *Brucella abortus*
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 07-Dec-1999
C:Accession: T31419
R:Inon de Iannino, N.; Briones, G.; Tolmasky, M.; Ugalde, R.A.
J. Bacteriol. 180, 4392-4400, 1998
A:Title: Molecular cloning and characterization of cgs, the *Brucella abortus* cyclic beta
chvB mutants.
A:Reference number: Z21023; MUID:98389650
A:Accession: T31419
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2831 <INO>
A:Cross-references: EMBL:AF047823; NID:g3551790; PID:g3551791; PIDN:AAC34747.1
A:Experimental source: strain S19
C:Genetics:
A:Gene: cgs

Query Match 89.7%; Score 26; DB 2; Length 2831;
Best Local Similarity 83.3%; Pred. No. 1.5e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
|||||
Db 1653 TAGDIK 1658

RESULT 14
T15789
hypothetical protein C41A3.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15789
R:Bentley, D.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of *C. elegans* cosmid C41A3.
A:Reference number: Z18404
A:Accession: T15789
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-7829 <BEN>
A:Cross-references: EMBL:U41541; NID:g1109867; PID:g1109869; PIDN:AAA83181.1; CESP:C4
C:Genetics:
A:Gene: CESP:C41A3.1
A:Introns: 29/2; 69/3; 141/2; 192/2; 271/3; 321/2; 1000/1; 1098/1; 1127/3; 1168/2; 12
/2; 7514/3; 7596/3; 7654/3; 7716/3; 7732/3; 7769/2; 7797/3
Query Match 89.7%; Score 26; DB 2; Length 7829;
Best Local Similarity 83.3%; Pred. No. 4.1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
|||||
Db 6807 SSGDIK 6812

RESULT 15
C69698
ribosomal protein L35 rpmI - *Bacillus subtilis*
C:Species: *Bacillus subtilis*
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: C69698
R:Kunst, F.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Bruchi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*
A:Reference number: A69580; MUID:98044033
A:Accession: C69698
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-66 <KUN>
A:Cross-references: GB:Z99118; GB:AL009126; NID:g2635200; PIDN:CAB14846.1; PID:g26353
A:Experimental source: strain 168
C:Genetics:
A:Gene: rpmI
C:Superfamily: *Escherichia coli* ribosomal protein L35

Query Match 86.2%; Score 25; DB 2; Length 66;
Best Local Similarity 83.3%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SAGDIK 6
|||||
Db 51 SAGDFK 56

Search completed: November 8, 2000, 08:53:28
Job time: 359 sec

A:Reference number: A75250; MUID:20036896
A:Accession: A75521
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-253 <WHI>
A:Cross-references: GB:AE001902; GB:AE000513; NID:g6458103; PIDN:AAF10004.1; PID:g645810
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0427
A:Map position: 1

Query Match 89.7%; Score 26; DB 2; Length 253;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
:|||||
DB 164 TAGDIK 169

RESULT 8
D69310
Grp cyclohydrolase II (ribA-1) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 16-Jul-1999
C:Accession: D69310
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Ulterback, T.; Cotton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343
A:Accession: D69310
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-378 <KLE>
A:Cross-references: GB:AE001070; GB:AE000782; NID:g2689393; PIDN:AA890751.1; PID:g265014
C:Superfamily: riba bifunctional protein; 3,4-dihydroxy-2-butanone 4-phosphate synthase
F:4-178/Domain: 3,4-dihydroxy-2-butanone 4-phosphate synthase homology <HBPS>
F:225-374/Domain: cyclohydrolase homology <CYCH>

Query Match 89.7%; Score 26; DB 2; Length 378;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
:|||||
DB 96 NAGDIK 101

RESULT 9
S55091
probable membrane protein YMR209c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YMR261.03c
C:Species: Saccharomyces cerevisiae
C:Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 05-Nov-1999
C:Accession: S55091
R:Dedman, K.; Brown, D.; Bowman, S.
submitted to the EMBL Data Library, June 1995
A:Reference number: S55089
A:Accession: S55091
A:Molecule type: DNA
A:Residues: 1-457 <DEP>
A:Cross-references: EMBL:Z49809; NID:g854459; PID:g854461; GSPDB:GN00013; MIPS:YMR209c
A:Experimental source: strain AB972
C:Genetics:
A:Gene: MIPS:YMR209c
A:Map position: 13R
C:Keywords: transmembrane protein

F:5-21/Domain: transmembrane #status predicted <TMM>

Query Match 89.7%; Score 26; DB 2; Length 457;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
:|||||
DB 242 SAGDMK 247

RESULT 10
S33907
glycerol kinase (EC 2.7.1.30) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YHL032c
C:Species: Saccharomyces cerevisiae
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 18-Jun-1999
C:Accession: S33907; S48936
R:Pavlik, P.; Simon, M.; Schuster, T.; Ruis, H.
Curr. Genet. 24, 21-25, 1993
A:Title: The glycerol kinase (GUT1) gene of Saccharomyces cerevisiae: cloning and cha
A:Reference number: S33907; MUID:93365032
A:Accession: S33907
A:Molecule type: DNA
A:Residues: 1-709 <PAV>
A:Cross-references: EMBL:X69049; NID:g312422; PIDN:CAA48791.1; PID:g312423
R:Favellio, T.
submitted to the EMBL Data Library, June 1994
A:Description: The sequence of S. cerevisiae cosmid 9196.
A:Reference number: S46794
A:Accession: S48936
A:Molecule type: DNA
A:Residues: 1-709 <FAV>
A:Cross-references: EMBL:U11583; NID:g2289854; PIDN:AA865044.1; PID:g2289861; MIPS:YH
C:Genetics:
A:Gene: SGD:GUT1
A:Cross-references: SGD:S0001024; MIPS:YHL032c
A:Map position: 8L
C:Superfamily: xylulokinase
C:Keywords: glycerol metabolism; phosphotransferase

Query Match 89.7%; Score 26; DB 2; Length 709;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
:|||||
DB 105 NAGDIK 110

RESULT 11
T13593
hypothetical protein 66A1.3 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jun-2000
C:Accession: T13593
R:Ferraz, C.; Vidal, S.; Brun, C.; Bucheton, A.; Demaille, J.G.
submitted to the EMBL Data Library, October 1998
A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A:Reference number: Z17692
A:Accession: T13593
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-958 <FER>
A:Cross-references: EMBL:AL031227; NID:e1330103; PID:e1330104; PIDN:CAA20226.1
C:Genetics:
A:Introns: 383/1; 706/3; 906/1
A:Note: EG:66A1.3

Query Match 89.7%; Score 26; DB 2; Length 958;

Query Match 100.0%; Score 29; DB 2; Length 550;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6

Db 310 SAGDIK 315

RESULT 3

Tl4323
cysteine proteinase inhibitor, extracellular - carrot
N;Alternate names: extracellular insoluble cystatin
C;Species: Daucus carota (carrot)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Jun-2000
C;Accession: Tl4323
R;Ojima, A.; Shiota, H.; Higashi, K.; Kamada, H.; Shimma, Y.; Wada, M.; Satoh, S.
Plant Mol. Biol. 34, 99-109, 1997
A;Title: An extracellular insoluble inhibitor of cysteine proteinases in cell cultures
A;Reference number: 217982; MUID:97320465

A;Accession: Tl4323

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-133 <OJI>

A;Cross-references: EMBL:D85623

A;Experimental source: strain US-Harumakigosun

C;Genetics:

A;Gene: EICC

C;Superfamily: cystatin; cystatin homology

C;Keywords: cysteine proteinase inhibitor

Query Match 93.1%; Score 27; DB 2; Length 133;

Best Local Similarity 83.3%; Pred. No. 39;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6

Db 73 SAGDIK 78

RESULT 4

Tl3600
hypothetical protein 80H7.11 - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jun-2000
C;Accession: Tl3600
R;Benos, P.
submitted to the EMBL Data Library, April 1999
A;Description: Sequencing the distal X chromosome of Drosophila melanogaster.

A;Reference number: 217667

A;Accession: Tl3600

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-352 <BEN>

A;Cross-references: EMBL:AL031027; NID:el313443; PID:el427311; PIDN:CAB41537.1

C;Genetics:

A;Note: EG:80H7.11

Query Match

Best Local Similarity 93.1%; Score 27; DB 2; Length 352;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6

Db 99 SAGDIK 104

RESULT 5

S75555
probable general stress protein Ctc - Deinococcus radiodurans (strain R1)

acyl carrier protein - Odontella sinensis chloroplast

C;Species: chloroplast Odontella sinensis

C;Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 20-Jun-2000

C;Accession: S78295

R;Kowalik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.

Plant Mol. Biol. Rep. 13, 336-342, 1995

A;Title: The Chloroplast Genome of a Chlorophyll a+c- containing Alga, Odontella sine

A;Reference number: S78238

A;Accession: S78295

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-80 <KOW>

A;Cross-references: EMBL:D67753; NID:gl185127; PIDN:CAA91668.1; PID:gl185185

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995

C;Genetics:

A;Gene: acp

A;Genome: Chloroplast

C;Function:

A;Description: carrier of the growing fatty acid chain; growing fatty acid chain is c

A;Pathway: fatty acid biosynthesis

C;Superfamily: acyl carrier protein; acyl carrier protein homology

C;Keywords: carrier protein; chloroplast; fatty acid biosynthesis; phosphopantetheine

F;3-74/Domain: acyl carrier protein homology <ACP>

F;38/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match

Best Local Similarity 89.7%; Score 26; DB 2; Length 80;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6

Db 60 AAGDIK 65

RESULT 6

S00737
hypothetical protein 1 - Methanococcus thermolithotrophicus (fragment)

C;Species: Methanococcus thermolithotrophicus

C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 22-Oct-1999

C;Accession: S00737

R;Souillard, N.; Magot, M.; Possot, O.; Sibold, L.

J. Mol. Evol. 27, 65-76, 1988

A;Title: Nucleotide sequence of regions homologous to nifH (nitrogenase Fe protein) f

ary implications.

A;Reference number: S00737; MUID:88259240

A;Accession: S00737

A;Molecule type: DNA

A;Residues: 1-162 <SOU>

A;Cross-references: EMBL:X07500; NID:g44619; PIDN:CAA30380.1; PID:g44620

A;Note: the authors translated the codon ATA for residue 146 as Leu

Query Match

Best Local Similarity 89.7%; Score 26; DB 2; Length 162;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6

Db 72 NAGDIK 77

RESULT 7

A75521

probable general stress protein Ctc - Deinococcus radiodurans (strain R1)

C;Species: Deinococcus radiodurans

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C;Accession: A75521

R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J

, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:53:26 ; Search time 99.87 Seconds
(without alignments)
3.813 Million cell updates/sec

Title: US-09-236-468a-2_COPY_310_315

Perfect score: 29

Sequence: 1 SAGDIK 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 182106

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_65.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	307	2 T04713	probable formamida
2	29	100.0	550	2 A57519	parathyroid hormon
3	27	93.1	133	2 T14323	cysteine proteinas
4	27	93.1	352	2 T13600	hypothetical prote
5	26	89.7	80	2 S78295	acyl carrier prote
6	26	89.7	162	2 S00737	hypothetical prote
7	26	89.7	253	2 A75521	probable general s
8	26	89.7	378	2 D69310	GTP cyclohydrolase
9	26	89.7	457	2 S55091	probable membrane
10	26	89.7	709	2 S33907	glycerol kinase (E
11	26	89.7	958	2 T13593	hypothetical prote
12	26	89.7	1168	2 S76195	hypothetical prote
13	26	89.7	2831	2 T31419	cyclic beta 1-2 gl
14	26	89.7	7829	2 T15789	hypothetical prote
15	25	86.2	66	2 C69698	ribosomal protein
16	25	86.2	88	2 A57711	diazepam-binding i
17	25	86.2	102	2 T32603	hypothetical prote
18	25	86.2	103	2 A81032	helix-turn-helix f
19	25	86.2	103	2 B81975	probable DNA-bindi
20	25	86.2	109	2 T34735	probable gas vesic
21	25	86.2	198	2 P00034	hypothetical prote
22	25	86.2	200	2 F64092	ribosomal protein
23	25	86.2	213	2 T06474	pore protein 24K c
24	25	86.2	214	2 T49758	hypoxanthine phosph
25	25	86.2	218	1 RTHUG	hypoxanthine phosph
26	25	86.2	218	1 RTMSC	hypoxanthine phosph
27	25	86.2	218	1 RTHYG	hypoxanthine phosph
28	25	86.2	218	2 S43043	hypoxanthine (guan
29	25	86.2	218	2 S21474	hypoxanthine phosph

30 25 86.2 218 2 I51842 hypoxanthine phosph
31 25 86.2 218 2 S18140 hypoxanthine phosph
32 25 86.2 240 2 T23797 hypothetical prote
33 25 86.2 247 2 S23474 rod-core linker po
34 25 86.2 258 2 T10083 expansin S2 precur
35 25 86.2 273 2 S73856 ornithine carbamoy
36 25 86.2 277 2 F71335 probable A/G-speci
37 25 86.2 316 2 E81845 thiorodoxin reduct
38 25 86.2 316 2 C81097 thiorodoxin reduct
39 25 86.2 328 2 H81229 DNA-directed RNA p
40 25 86.2 336 2 A72247 DNA-directed RNA p
41 25 86.2 340 2 D70318 phosphoribosylanth
42 25 86.2 342 2 JC4092 DMCI/LIM15 homolog
43 25 86.2 343 2 C75260 conserved hypothet
44 25 86.2 345 2 T08838 RecA/Rad51/DMC1-1i
45 25 86.2 349 2 JC2214 hypothetical 38.3K

ALIGNMENTS

RESULT 1

T04713

probable formamidase (EC 3.5.1.49) F19F18.50 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999

C:Accession: T04713

R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.

Submitted to the Protein Sequence Database, March 1999

A:Reference number: Z15382

A:Accession: T04713

A:Molecule type: DNA

A:Residues: 1-307 <BEV>

A:Cross-references: EMBL:AL035605

A:Experimental source: cultivar Columbia; BAC clone F19F18

C:Genetics:

A:Map position: 4

A:Introns: 73/3; 153/1; 186/2; 267/3

A:Note: F19F18.50

C:Keywords: hydrolase

Query Match 100.0%; Score 29; DB 2; Length 307;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6

DB 62 SAGDIK 67

RESULT 2

A57519

parathyroid hormone receptor 2 precursor - human

N:Alternate names: PTH2 receptor

C:Species: Homo sapiens (man)

C>Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 24-Nov-1999

C:Accession: A57519

R:Udolin, T.B.; Gruber, C.; Bonner, T.I.

J. Biol. Chem. 270, 15455-15458, 1995

A:Title: Identification and functional expression of a receptor selectively recognizi

A:Reference number: A57519; MUID:95318121

A:Accession: A57519

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-550 <USD>

A:Cross-references: GB:U25128; NID:g887966; PIDN:AAC50157.1; PID:g887967

C:Genetics:

A:Gene: GDB:PTH2; PTHR2

A:Cross-references: GDB:731977; OMIM:601469

A:Map position: 2q33-2q33

C:Superfamily: glucagon receptor

C:Keywords: hormone receptor

THIS PAGE BLANK (USPTO)

RESULT 14
YHGB_PSEPU
ID YHGB_PSEPU STANDARD; PRT; 198 AA.
AC P20162;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PROBABLE ABC TRANSPORTER ATP-BINDING PROTEIN IN NTRA/RPON 5'REGION
DE (ORF1) (FRAGMENT).
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TN2100;
RX MEDLINE; 90152355.
RA Inouye S., Yamada M., Nakazawa A., Nakazawa T.;
RT "Cloning and sequence analysis of the ntrA (rpoN) gene of Pseudomonas
putida.";
RL Gene 85:145-152(1989).
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
(ABC TRANSPORTERS). STRONG, TO E.COLI YHGB AND TO CORRESPONDING
PROTEIN IN OTHER BACTERIA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

CC EMBL; M24916; AAA88443.1; ALT_TERM.
DR PIR; PQ0034; PQ0034.
DR INTERPRO; IPR001617; -.
DR PFAM; PF00005; ABC_tran; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Transport.
FT NON_TER 1
SQ SEQUENCE 198 AA; 22272 MW; 2AFCC053AF958A22 CRC64;

Query Match 86.2%; Score 25; DB 1; Length 198;
Best Local Similarity 83.3%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 SAGDIK 6
I I I I I
Db 70 SVGDIK 75

RESULT 15
RL4_HAEIN
ID RL4_HAEIN STANDARD; PRT; 200 AA.
AC P44345;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L4.
GN RPLD OR RPLA OR HI0778.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20;
RX MEDLINE; 95350630.
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McInerney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,

Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
Fine L.D., Fritchman J.L., Fuhrmann J.L., Geohagen N.S.M.,
Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
influenzae Rd.";
RL Science 269:496-512(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Zengel J.M., Vorozheikina D., Li X., Lindahl L.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY AND SPECIFICALLY TO 23S
RNA (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE L4P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

CC EMBL; U32761; AAC22437.1; -.
DR EMBL; U37797; AAB41511.1; -.
DR TIGR; HI0778; -.
DR INTERPRO; IPR002136; -.
DR PFAM; PF00573; Ribosomal_L4; 1.
KW Ribosomal protein; rRNA-binding.
FT CONFLICT 65 65 G->A (IN REF. 2).
SQ SEQUENCE 200 AA; 21954 MW; A135A4762FE50927 CRC64;

Query Match 86.2%; Score 25; DB 1; Length 200;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 AGDIK 6
I I I I I
Db 69 AGDIK 73

Search completed: November 8, 2000, 09:03:49
Job time: 859 sec

RX MEDLINE; 92319720.
RA Malagon M., Vaudry H., Vallarino M., Gracia-Navarro F., Tonon M.C.;
RT "Distribution and characterization of endorepine-like
RT immunoreactivity in the central nervous system of the frog Rana
RT ridibunda.";
RL Peptides 13:99-107(1992).
CC -1- FUNCTION: MAY PLAY IMPORTANT FUNCTIONS IN THE CONTROL OF BRAIN AND
CC PITUITARY ACTIVITIES. MAY REGULATE GABA NEUROTRANSMISSION THROUGH
CC A PARACRINE AND/OR AUTOCRINE MECHANISM. MAY NOT BIND ACYL-COA
CC ESTERS.
CC -1- TISSUE SPECIFICITY: BRAIN. IS SELECTIVELY EXPRESSED IN GLIAL
CC CELLS.
CC -1- SIMILARITY: BELONGS TO THE ACPF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U09205; AAB60606.1; -.
CC HSSP; P07107; 2ABD.
CC INTERPRO; IPR000582; -.
CC PFAM; PF00887; ACPB; 1.
CC PRINTS; PR00689; ACPBINDINGP.
CC PROSITE; PS00880; ACPB; 1.
CC Transprot; Lipid-binding.
CC INIT_MET 0
CC PEPTIDE 18 51 TRIAKONTATETRAPEPTIDE (TTN)
CC
CC PEPTIDE 34 51 OCTADECANEUROPEPTIDE (ODN) (POTENTIAL).
CC SEQUENCE 87 AA; 9677 MW; F6621EFC12054BA CRC64;
CC
Query Match 86.2%; Score 25; DB 1; Length 87;
Best Local Similarity 66.7%; Pred. No. 35;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 SAGDIK 6
: : : : :
Db 9 AAGDVK 14
RESULT 12
Y113_METEX STANDARD; PRT; 88 AA.
AC 005113;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL 9.7 KDA PROTEIN (ORF1).
OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AM1 / NCIB 9133;
RX MEDLINE; 97386438.
RA Chistoserdova L., Lidstrom M.E.;
RT "Identification and mutation of a gene required for glycerate kinase
RT activity from a facultative methylophilic, Methylobacterium extorquens
RT AM1.";
RL J. Bacteriol. 179:4946-4948(1997).
CC -1- SIMILARITY: STRONG. TO R.PROWAZEKII RP113.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
CC or send an email to license@isb-sib.ch).
CC

CC or send an email to license@isb-sib.ch).
CC EMBL; U87316; AAB66493.1; -.
KW Hypothetical protein.
SQ SEQUENCE 88 AA; 9706 MW; C501A82AA35855F8 CRC64;
Query Match 86.2%; Score 25; DB 1; Length 88;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 AGDIK 6
: : : : :
Db 37 AGDIK 41
RESULT 13
TCP2_PIG STANDARD; PRT; 104 AA.
AC Q29236;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE T-COMPLEX PROTEIN 1, ZETA SUBUNIT (TCP-1-ZETA) (CCT-ZETA) (FRAGMENT).
GN CCT6 OR CCTZ.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;
RX MEDLINE; 96327607.
RA Winteroe A.K., Fredholm M., Davies W.;
RT "Evaluation and characterization of a porcine small intestine cDNA
RT library: analysis of 839 clones.";
RL Mamm. Genome 7:509-517(1996).
CC -1- FUNCTION: MOLECULAR CHAPERONE; ASSIST THE FOLDING OF PROTEINS UPON
CC ATP HYDROLYSIS. KNOWN TO PLAY A ROLE, IN VITRO, IN THE FOLDING OF
CC ACTIN AND TUBULIN (BY SIMILARITY).
CC -1- SUBUNIT: HETERO-OLIGOMERIC COMPLEX OF ABOUT 850 TO 900 KDA THAT
CC FORMS TWO STACKED RINGS, 12 TO 16 NM IN DIAMETER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE TCP-1 CHAPERONIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; F14833; CAA23284.1; -.
CC INTERPRO; IPR002194; -.
CC INTERPRO; IPR002423; -.
CC PFAM; PF00118; cpn60_TCP1; 1.
CC PROSITE; PS00750; TCP1_1; 1.
CC PROSITE; PS00751; TCP1_2; 1.
CC PROSITE; PS00995; TCP1_3; PARTIAL.
KW Chaperone; ATP-binding; Multigene family.
FT NON_TER 1
FT NON_TER 104 104
SQ SEQUENCE 104 AA; 10937 MW; AFB1758393FFA606 CRC64;
Query Match 86.2%; Score 25; DB 1; Length 104;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 AGDIK 6
: : : : :
Db 36 AGDIK 40

CC -1- FUNCTION: TRANSCRIPTION FACTOR, MASTER REGULATOR OF TRACHEAL CELL
CC FATES IN THE EMBRYO, NECESSARY FOR THE DEVELOPMENT OF THE SALIVARY
CC GLAND DUCT AND THE POSTERIOR SPIRACLES. IT MAY INDUCE A GENERAL
CC FATE OF BRANCHED TUBULAR STRUCTURES OF EPITHELIAL ORIGIN. TGO/TRH
CC HETERODIMERS ARE INVOLVED IN THE CONTROL OF BREATHLESS EXPRESSION.
CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC BHLH PROTEIN. HETERODIMER WITH TGO.
CC -1- SUBCELLULAR LOCATION: NUCLEAR PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1 (SHOWN HERE), 2 AND
CC 3; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: TRACHEA, SALIVARY GLAND DUCTS, POSTERIOR
CC SPIRACLES (FILZKOEPER PRIMORDIA) AND A SUBSET OF CELLS IN THE CNS.
CC -1- DEVELOPMENTAL STAGE: DURING EMBRYOGENESIS, FIRST DETECTED IN THE
CC TRACHEAL PLACODES AT STAGE 8, AND EXPRESSION CONTINUES THROUGHOUT
CC EMBRYONIC AND LARVAL DEVELOPMENT. IN THE DEVELOPING SALIVARY
CC GLAND, EXPRESSION IS OBSERVED IN THE ENTIRE GLAND AT STAGE 9 AND
CC BY STAGE 12, EXPRESSION IS CONFINED TO THE SALIVARY DUCTS.
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U33427; AAA96257.1; -;
DR EMBL; U42699; AAA96754.1; -;
DR FLYBASE; FBgn0003749; trh.
DR INTERPRO; IPR000014; -;
DR INTERPRO; IPR003015; -;
DR PFAM; PF00989; PAS; 2.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
KW Developmental protein; Nuclear protein; Transcription regulation;
KW Repeat; DNA-binding; Alternative splicing.
FT DNA_BIND 77 90
FT DOMAIN 91 131
FT REPEAT 167 233
FT REPEAT 384 450
FT DOMAIN 456 499
FT DOMAIN 620 627
FT DOMAIN 145 148
FT DOMAIN 235 241
FT DOMAIN 242 245
FT DOMAIN 246 255
FT DOMAIN 713 719
FT VARSPPLIC 272 277
FT VARSPPLIC 319 347
FT CONFLICT 69 69
FT CONFLICT 241 241
FT CONFLICT 694 694
FT CONFLICT 699 699
FT CONFLICT 820 820
SQ SEQUENCE 949 AA; 101354 MW; 12C954F868CF9E1A CRC64;

Query Match 89.7%; Score 26; DB 1; Length 949;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6
Db 546 SAGDMK 551

RESULT 10
RL35_BACSU STANDARD; PRT; 65 AA.
AC P55874;
DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L35.
GN RPMI.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RX MEDLINE; 97124191.
RA Wipat A., Carter N., Brignell C.S., Guy J.B., Piper K.,
RA Sanders J., Emerson P.T., Harwood C.R.;
RT "The dnaB-phoA (256 degrees-240 degrees) region of the Bacillus
RT subtilis chromosome containing genes responsible for stress
RT responses, the utilization of plant cell walls and primary
RT metabolism";
RL Microbiology 142:3067-3078(1996).
CC -1- SIMILARITY: BELONGS TO THE L35P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z75208; CAA99617.1; -;
DR EMBL; Z99118; CAB14846.1; -;
DR SUBTILIST; BG11972; RPMI.
DR INTERPRO; IPR001706; -;
DR PFAM; PF01632; Ribosomal_L35p; 1.
DR PRINTS; PR00064; RIBOSOMAL_L35.
DR PROSITE; PS00936; RIBOSOMAL_L35; 1.
KW Ribosomal protein.
FT INIT_MET 0
FT INIT_MET 0 BY SIMILARITY.
SQ SEQUENCE 65 AA; 7426 MW; 06FE064FAB3F30B CRC64;

Query Match 86.2%; Score 25; DB 1; Length 65;
Best Local Similarity 83.3%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SAGDIK 6
Db 50 SAGDFK 55

RESULT 11
ACBP_RANRI STANDARD; PRT; 87 AA.
ID ACBP_RANRI
AC P45883;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE ACYL-COA-BINDING PROTEIN HOMOLOG (ACBP) (DIAZEPAM BINDING INHIBITOR
DE HOMOLOG) (DB1).
OS Rana ridibunda (laughing frog) (Marsh frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-39 AND 58-87.
RC TISSUE-BRAIN;
RX MEDLINE; 94316605.
RA Lihrmann I., Plaquent J.-C., Tostivint H., Rajmakers R.,
RA Tonon M.-C., Conlon J.M., Vaudry H.;
RT "Frog diazepam-binding inhibitor: peptide sequence, cDNA cloning, and
RT expression in the brain";
RL Proc. Natl. Acad. Sci. U.S.A. 91:6899-6903(1994).
RN [2]
RP CHARACTERIZATION.
RC TISSUE-BRAIN;

RESULT 7
 YM59_YEAST STANDARD; PRT; 457 AA.
 AC Q03648;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE HYPOTHETICAL 52.2 KDA PROTEIN IN RARI-SCJ1 INTERGENIC REGION.
 GN YMR209C OR YMR261.03C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomycetes.
 RN [1]
 RC STRAIN=S288C / AB972;
 RA Dedman K., Brown D., Bowman S., Barrell B.G., Rajandream M.A.,
 RA Walsh S.V.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z49809; CAA89924.1; -
 DR SGD: S0004822; YMR209C.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 1 21 POTENTIAL.
 FT TRANSMEM 250 270 POTENTIAL.
 SQ SEQUENCE 457 AA; 52246 MW; B68737D1E58176E3 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 457;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
 DB 242 SAGDMK 247
 |||||

RESULT 8
 GLPK_YEAST STANDARD; PRT; 709 AA.
 AC P32190;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE GLYCEROL KINASE (EC 2.7.1.30) (ATP:GLYCEROL 3-PHOSPHOTRANSFERASE)
 DE (GLYCEROKINASE) (GK).
 GN GUT1 OR YHL032C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomycetes.
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE; 93365032.
 RA Pavlik P., Simon M., Schuster T., Ruis H.;
 RT "The glycerol kinase (GUT1) gene of Saccharomyces cerevisiae: cloning
 RT and characterization.";
 RT Curr. Genet. 24:21-25(1993).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE; 94378003.
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 RA Du Z., Favella A., Fulton L., Gattung S., Geisel C., Kirsten J.,

RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
 RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
 RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
 RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
 RA Vaudin M.;
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 RT VIII.";
 RL Science 265:2077-2082(1994).
 CC -1- CATALYTIC ACTIVITY: ATP + GLYCEROL = ADP + GLYCEROL 3-PHOSPHATE.
 CC -1- PATHWAY: RATE-LIMITING STEP IN GLYCEROL UTILIZATION.
 CC -1- SIMILARITY: BELONGS TO THE FUCOKINASE / GLUCOKINASE /
 CC GLYCEROKINASE / XYLOKINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X69049; CAA48791.1; -
 DR EMBL: U11583; AAB65044.1; -
 DR PIR: S33907; S33907.
 DR PIR: S48936; S48936.
 DR HSP: P08859; IGLB.
 DR SGD: S001024; GUT1.
 DR INTERPRO: IPR000577; -
 DR PRAM: PF00370; FGGY; 4.
 DR PROSITE: PS00445; FGGY_KINASES_2; 1.
 DR PROSITE: PS00933; FGGY_KINASES_1; 1.
 KW Glycerol metabolism; Transferase; Kinase.
 SQ SEQUENCE 709 AA; 79824 MW; 12B42C4DABE49FF6 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 709;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
 DB 105 NAGDIK 110
 :|||||

RESULT 9
 TRH_DROME STANDARD; PRT; 949 AA.
 AC Q24119; Q24165;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE TRACHEALESS PROTEIN.
 GN TRH.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 RN [1]
 RC SEQUENCE FROM N.A.
 RC TISSUE=EMBRYO;
 RX MEDLINE; 96136712.
 RA Wilk R., Weizman I., Shilo B.-2.;
 RT "Tracheless encodes a bHLH-PAS protein that is an inducer of
 RT tracheal cell fates in Drosophila.";
 RL Genes Dev. 10:93-102(1996).
 RN [2]
 RP SEQUENCE OF 27-949 FROM N.A.
 RC TISSUE=EMBRYO;
 RX MEDLINE; 96136713.
 RA Isaac D.D., Andrew D.J.;
 RT "Tubulogenesis in Drosophila: a requirement for the tracheless gene
 RT product.";
 RL Genes Dev. 10:103-117(1996).

```
CC -!- SIMILARITY: TO BACTERIAL AND TO PLANT NUCLEAR-ENCODED ACYL
CC CARRIER PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 267753; CAA91668.1; -.
CC HSSP; P02901; IACP.
CC MENDEL; 13226; ODOsi; Ac11.1.
CC INTERPRO; IPR002025; -.
CC PFAM; PF00550; pp-binding; 1.
CC PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
CC PROSITE; PS00073; ACP_DOMAIN; 1.
CC Fatty acid biosynthesis; Phosphopantetheine; Chloroplast.
FT BINDING 38 38 PHOSPHOPANTHETHEINE (BY SIMILARITY).
SQ SEQUENCE 80 AA; 8945 MW; 5AF2E2590A7412AE CRC64;

Query Match 89.7%; Score 26; DB 1; Length 80;
Best Local Similarity 83.3%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
DB 60 AAGDIK 65

RESULT 5
YNIL_METTL STANDARD; PRT; 162 AA.
ID YNII_METTL
AC P03409;
RA Souillard N., Magot M., Possot O., Sibold L.;
RT "Nucleotide sequence of regions homologous to nifH (nitrogenase Fe
protein) from the nitrogen-fixing archaeobacteria Methanococcus
thermolithotrophicus and Methanobacterium ivanovii: evolutionary
implications.";
RL J. Mol. Evol. 27:65-76(1988).
CC -!- SIMILARITY: BELONGS TO THE UPF0020 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X07500; CAA30380.1; -.
CC PIR; S00737; S00737.
CC HSSP; P14385; IAQI.
CC INTERPRO; IPR000241; -.
CC PROSITE; IPR002052; -.
CC PROSITE; PS01261; UPF0020; 1.
CC PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 162 AA; 18186 MW; 08EE977BE7F5622E CRC64;
```

```
Query Match 89.7%; Score 26; DB 1; Length 162;
Best Local Similarity 83.3%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
DB 72 NAGDIK 77

RESULT 6
GCH2_ARCFU STANDARD; PRT; 378 AA.
ID GCH2_ARCFU
AC O29766;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GTP CYCLOHYDROLASE II (EC 3.5.4.25).
GN RIBA OR AF0484.
OS Archaeoglobus fulgidus.
CC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
CC Archaeoglobus.
CC [1]
RN SEQUENCE FROM N.A.
RP STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE; 98049343.
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Keriavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -!- CATALYTIC ACTIVITY: GTP + 3 H(2)O = FORMATE + 2,5-DIAMINO-6-
HYDROXY-4-(5-PHOSPHORIBOSYLAMINO)PYRIMIDINE + PYROPHOSPHATE.
CC -!- PATHWAY: RIBOFLAVIN BIOSYNTHESIS.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE DHBP
SYNTHASE FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE GTP
CYCLOHYDROLASE II FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001070; AAB90751.1; -.
CC TIGR; AF0484; -.
CC PFAM; PF00925; GTP_cyclohydro2; 1.
CC PFAM; PF00926; DHBp_synthase; 1.
CC Riboflavin biosynthesis; Hydrolase.
KW DOMAIN 1 180 DHBp SYNTHASE-LIKE.
FT DOMAIN 181 378 GTP CYCLOHYDROLASE II.
SQ SEQUENCE 378 AA; 42271 MW; 5F63B954488173B6 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 378;
Best Local Similarity 83.3%; Pred. No. 92;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
DB 96 NAGDIK 101
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=BRAIN;
RA MEDLINE: 95318121.
RA Usdin T.B., Gruber C., Bonner T.I.;
RT "Identification and functional expression of a receptor selectively
RT recognizing parathyroid hormone, the PTH2 receptor.";
RL J. Biol. Chem. 270:15455-15458(1995).
RN [2]
RP SEQUENCE OF 26-40 AND 306-550 FROM N.A.
RX MEDLINE: 97079671.
RA Usdin T.B., Modi W., Bonner T.I.;
RT "Assignment of the human PTH2 receptor gene (PTH2) to chromosome 2q33
RT by fluorescence in situ hybridization.";
RL Genomics 37:140-141(1996).
CC -1- FUNCTION: THIS IS A SPECIFIC RECEPTOR FOR PARATHYROID HORMONE. THE
CC ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLATE CYCLASE.
CC -1- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN BRAIN AND PANCREAS.
CC ALSO EXPRESSED IN THE TESTIS.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U25128; AAC50157.1; -;
DR EMBL: U47124; AAA96796.1; -;
DR EMBL: U47129; AAC50767.1; -;
DR EMBL: U47125; AAC50767.1; JOINED.
DR EMBL: U47126; AAC50767.1; JOINED.
DR EMBL: U47127; AAC50767.1; JOINED.
DR EMBL: U47128; AAC50767.1; JOINED.
DR GCRDB: GCR_2003; -;
DR MIM: 601469; -;
DR INTERPRO: IPR000832; -;
DR PFAM: PF00002; 7um2; 1.
DR PRINTS: PR00249; GPCRSECRETIN.
DR PROSITE: PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE: PS00650; G_PROTEIN_RECEP_F2_2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 550 PARATHYROID HORMONE RECEPTOR.
FT DOMAIN 27 145 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 146 169 1 (POTENTIAL).
FT DOMAIN 170 176 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 177 196 2 (POTENTIAL).
FT DOMAIN 197 237 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 238 260 3 (POTENTIAL).
FT DOMAIN 261 275 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 276 297 4 (POTENTIAL).
FT DOMAIN 298 316 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 317 337 5 (POTENTIAL).
FT DOMAIN 338 364 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 365 383 6 (POTENTIAL).
FT DOMAIN 384 394 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 395 417 7 (POTENTIAL).
FT DOMAIN 418 550 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 550 AA; 62235 MW; 2ADD14DBA68A9BF8 CRC64;

Query Match 100.0%; Score 29; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 27;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SAGDIK 6
Db 310 SAGDIK 315
RESULT 3
PRVB_LEUCE ID PRVB_LEUCE STANDARD; PRT; 106 AA.
AC P05939;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-NOV-1988 (Rel. 09, Last annotation update)
DE PARVALBUMIN BETA (V).
OS Leuciscus cephalus (Chub).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Leuciscinae; Leuciscus.
RN [1]
RP SEQUENCE.
RA Gerday C., Collins S., Piront A.;
RT "Phylogenetic relationships between Cyprinidae parvalbumins-II. The
RT amino acid sequence of the Parvalbumin V of Chub (Leuciscus Cephalus
RT L.).";
RL Comp. Biochem. Physiol. 61B:451-457(1978).
CC -1- FUNCTION: IN MUSCLE, THE CALCIUM-BINDING PROTEIN PARVALBUMIN IS
CC THOUGHT TO BE INVOLVED IN MUSCLE RELAXATION.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO
CC THE PARVALBUMINS SUBFAMILY.
DR HSSP: P02618; 4CPV.
DR INTERPRO: IPR002048; -;
DR PFAM: PF00036; efhand; 2.
DR PROSITE: PS00018; EF_HAND; 2.
KW Calcium-binding; Muscle protein; Duplication; Acetylation.
FT MOD_RES 1 1 ACETYLATION.
FT CA_BIND 51 62 SITE 1.
FT CA_BIND 90 101 SITE 2.
SQ SEQUENCE 106 AA; 11263 MW; BF27C2A24AFCB4D CRC64;
Query Match 96.6%; Score 28; DB 1; Length 106;
Best Local Similarity 83.3%; Pred. No. 8.4;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SAGDIK 6
Db 37 SAGDVK 42
RESULT 4
ACP_ODOSI ID ACP_ODOSI STANDARD; PRT; 80 AA.
AC P49517;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ACYL CARRIER PROTEIN.
GN ACPP OR ACIL OR ACP.
OS Odontella sinensis.
OG Chloroplast.
OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC Biddulphiophycidae; Eupodiscales; Eupodiscaeae; Odontella.
RN [1]
RP SEQUENCE FROM N.A.
RA Kowalik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
RT "The chloroplast genome of a chlorophyll a+c-containing alga,
RT Odontella sinensis.";
RL Plant Mol. Biol. Rep. 13:336-342(1995).
CC -1- FUNCTION: THIS PROTEIN IS THE CARRIER OF THE GROWING FATTY ACID
CC CHAIN IN FATTY ACID BIOSYNTHESIS.
CC -1- PTM: THE GROWING FATTY ACID CHAINS ARE COVALENTLY BOUND TO THE
CC 4'-PHOSPHOPANTHETINE PROSTHETIC GROUP (BY SIMILARITY).

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 09:03:48 ; Search time 58.45 Seconds
(without alignments)
3.280 Million cell updates/sec

Title: US-09-236-468a-2_copy_310_315

Perfect score: 29

Sequence: 1 SAGDIK 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 87993

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	29	100.0	116	1 RLA2_BRAFL	O01725 branchiostoma
2	29	100.0	550	1 PTR2_HUMAN	P49190 homo sapien
3	28	96.6	106	1 PRVB_LEUCE	P05939 leuciscus c
4	26	89.7	80	1 ACP_ODOSI	P49517 odontella s
5	26	89.7	162	1 YN11_METTL	P05409 methanococc
6	26	89.7	378	1 GCH2_ARCFU	O29766 methanoglob
7	26	89.7	457	1 YMS9_YEAST	Q03648 saccharomyc
8	26	89.7	709	1 GLPK_YEAST	P32190 saccharomyc
9	26	89.7	949	1 TRH_DROME	Q24119 drosophila
10	25	86.2	65	1 RL35_BACSU	P55874 bacillus su
11	25	86.2	87	1 ACBP_RANRI	P45883 rana ridibu
12	25	86.2	88	1 Y113_METEX	O05113 methyllobact
13	25	86.2	104	1 TCP2_PIG	Q29236 sus scrofa
14	25	86.2	198	1 YHBG_PSEPU	P20162 pseudomonas
15	25	86.2	200	1 RL4_HAEIN	P44345 haemophilus
16	25	86.2	213	1 HPRT_MUSSP	Q64531 mus spretus
17	25	86.2	217	1 HPRT_CRIGR	P00494 cricetus
18	25	86.2	217	1 HPRT_HUMAN	P00492 homo sapien
19	25	86.2	217	1 HPRT_MERUN	P47959 meriones un
20	25	86.2	217	1 HPRT_MOUSE	P00493 mus musculus
21	25	86.2	218	1 HPRT_RAT	P27605 rattus norv
22	25	86.2	247	1 PVG2_MASLA	P29732 mastigoclad
23	25	86.2	263	1 CBLJ_SALTU	Q05591 salmonella
24	25	86.2	273	1 OTCC_MYCPN	P75473 mycoplasma
25	25	86.2	336	1 RPOA_THEMEA	Q9x112 thermotoga
26	25	86.2	340	1 TRPD_AQUAE	O66576 aquifex aeo
27	25	86.2	342	1 DMCL_ARATH	Q39009 arabidopsis
28	25	86.2	345	1 DMCL_SOYBN	Q96449 glycine max
29	25	86.2	349	1 DMCL_LILLO	P37384 lilium long
30	25	86.2	412	1 TGF2_CHICK	P30371 gallus gall
31	25	86.2	418	1 YJEH_ECOLI	P39277 escherichia
32	25	86.2	445	1 HGD_HUMAN	Q93099 homo sapien
33	25	86.2	445	1 HGD_MOUSE	O09173 mus musculus

```
34 25 86.2 445 1 HPPD_ARATH      P93836 arabidopsis
35 25 86.2 513 1 Y4WI_RHISN      P55687 rhizobium s
36 25 86.2 514 1 VLL_HPVA7       P22424 human papil
37 25 86.2 520 1 C11A_CAPHI      P79153 capra hircu
38 25 86.2 531 1 TCPY_MOUSE      Q61390 mus musculu
39 25 86.2 531 1 TCP2_HUMAN      P40227 homo sapien
40 25 86.2 531 1 TCP2_MOUSE      P80317 mus musculu
41 25 86.2 531 1 TCP2_RABIT      O77622 oryctolagus
42 25 86.2 538 1 NRFA_HAEIN      P45017 haemophilus
43 25 86.2 539 1 TCP2_CAEEL      P46550 caenorhabdi
44 25 86.2 540 1 TCPW_HUMAN      Q92526 homo sapien
45 25 86.2 550 1 VGLE_HSV11      P04488 herpes simp
```

ALIGNMENTS

```
RESULT 1
RLA2_BRAFL
ID RLA2_BRAFL STANDARD; PRT; 116 AA.
AC O01725;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 60S ACIDIC RIBOSOMAL PROTEIN P2.
OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
RN [1]
RP, SEQUENCE FROM N.A.
RA Tweedie S., Charlton J., Clark V., Bird A.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
CC PROTEIN SYNTHESIS.
CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
CC SUBUNIT.
CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z83263; CAB05855.1; -.
DR INTERPRO; IPR001813; -.
DR PFAM; PF00428; 60s.Ribosomal; 1.
KW Ribosomal protein; Phosphorylation.
SQ SEQUENCE 116 AA; 11557 MW; 39C16854F10DC00C CRC64;

Query Match 100.0%; Score 29; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6
Db 19 SAGDIK 24

RESULT 2
PTR2_HUMAN
ID PTR2_HUMAN STANDARD; PRT; 550 AA.
AC P49190;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE PARATHYROID HORMONE RECEPTOR PRECURSOR (PTH2 RECEPTOR).
GN PTH2.
OS Homo sapiens (Human).
```


Thu Nov 9 08:56:44 2000

us-09-236-468a-2_copy_310_315.ra1

Page 8

Search completed: November 8, 2000, 08:49:19
Job time: 112 sec

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note="S0.50;L0.50"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 3
; OTHER INFORMATION: /note="T0.34;E0.33;K0.33"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4
; OTHER INFORMATION: /note="A0.34;T0.33;P0.33"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 5
; OTHER INFORMATION: /note="I0.50;V0.50"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 6
; OTHER INFORMATION: /note="F0.50;L0.50"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 8
; OTHER INFORMATION: /note="T0.50;V0.50"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 18
; OTHER INFORMATION: /note="A0.50;K0.50"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 20
; OTHER INFORMATION: /note="D0.34;T0.33;E0.33"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 22
; OTHER INFORMATION: /note="K0.50;V0.50"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 23
; OTHER INFORMATION: /note="T0.34;A0.33;S0.33"
```

```
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 24
; OTHER INFORMATION: /note="S0.34;G0.33;N0.33"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 27
; OTHER INFORMATION: /note="G0.50;N0.50"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 28
; OTHER INFORMATION: /note="Q0.50;E0.50"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 30
; OTHER INFORMATION: /note="G0.50;A0.50"
; US-08-488-351A-86

Query Match      82.8%; Score 24; DB 2; Length 30;
Best Local Similarity 80.0%; Pred. No. 26;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AGDIK 6
Db 18 AGDVK 22

RESULT 15
PCT-US93-08739-3
; Sequence 3, Application PC/TUS9308739
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented
; TITLE OF INVENTION: SYNTHETIC PEPTIDE VACCINE FOR
; TITLE OF INVENTION: CHLAMYDIA
; TITLE OF INVENTION: TRACHOMATIS
; NUMBER OF SEQUENCES: 3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08739
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/947,671 US
; FILING DATE: 18 SEP 93
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; STRANDEDNESS: Single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; PCT-US93-08739-3

Query Match      82.8%; Score 24; DB 4; Length 42;
Best Local Similarity 80.0%; Pred. No. 37;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AGDIK 6
Db 38 AGDVK 42
```

SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-374-560-4

Query Match 82.8%; Score 24; DB 2; Length 19;
Best Local Similarity 80.0%; Pred. No. 16;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGDIK 6
111:1
Db 10 AGDVK 14

RESULT 13

US-08-446-692-86
Sequence 86, Application US/08446692
Patent No. 5759551
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "S0.50;L0.50"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 3
OTHER INFORMATION: /note= "T0.34;E0.33;K0.33"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 4
OTHER INFORMATION: /note= "A0.34;T0.33;P0.33"
FEATURE:
NAME/KEY: Modified-site

LOCATION: 5
OTHER INFORMATION: /note= "I0.50;V0.50"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 6
OTHER INFORMATION: /note= "F0.50;L0.50"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 8
OTHER INFORMATION: /note= "T0.50;V0.50"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 18
OTHER INFORMATION: /note= "A0.50;K0.50"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 20
OTHER INFORMATION: /note= "D0.34;T0.33;E0.33"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 22
OTHER INFORMATION: /note= "K0.50;V0.50"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 23
OTHER INFORMATION: /note= "T0.34;A0.33;S0.33"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 24
OTHER INFORMATION: /note= "S0.34;G0.33;N0.33"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 27
OTHER INFORMATION: /note= "G0.50;N0.50"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 28
OTHER INFORMATION: /note= "Q0.50;E0.50"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 30
OTHER INFORMATION: /note= "G0.50;A0.50"
US-08-446-692-86

Query Match 82.8%; Score 24; DB 1; Length 30;
Best Local Similarity 80.0%; Pred. No. 26;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGDIK 6
111:1
Db 18 AGDVK 22

RESULT 14

US-08-488-351A-86
Sequence 86, Application US/08488351A
Patent No. 5843446
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:

```
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 2 AGDIK 6
   |||:|
Db 5 AGDVK 9

RESULT 10
; Sequence 67, Application US/08485286
; Patent No. 5646026
; Patent No. 5646026 5646119
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485.286
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/378761
; FILING DATE: 26-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-286-67

Query Match 82.8%; Score 24; DB 1; Length 11;
Best Local Similarity 80.0%; Pred. No. 9;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 2 AGDIK 6
   |||:|
Db 5 AGDVK 9

RESULT 11
PCT-US93-08739-2
; Sequence 2, Application PC/TUS9308739
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented
; APPLICANT: by the Secretary of the Department of Health and Human Services
; TITLE OF INVENTION: SYNTHETIC PEPTIDE VACCINE FOR
; TITLE OF INVENTION: CHLAMYDIA

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 2 AGDIK 6
   |||:|
Db 5 AGDVK 9

RESULT 12
US-08-374-560-4
; Sequence 4, Application US/08374560
; Patent No. 5882645
; GENERAL INFORMATION:
; APPLICANT: TOTH, Islván
; APPLICANT: GIBBONS, William Anthony
; TITLE OF INVENTION: PEPTIDE COMPOUNDS
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
; STREET: 1180 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10036-8403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,560
; FILING DATE: 13-MAR-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9215780.9
; FILING DATE: 24-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Meilman, Edward A.
; REGISTRATION NUMBER: 24,735
; REFERENCE/DOCKET NUMBER: P/365-302
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 382-0700
; TELEFAX: (212) 382-0888
; TELEX: 236925
; INFORMATION FOR SEQ ID NO: 4:
```

```
; TITLE OF INVENTION: TRACHOMATIS
; NUMBER OF SEQUENCES: 3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08739
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/947,671 US
; FILING DATE: 18 SEP 93
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; PCT-US93-08739-2

Query Match 82.8%; Score 24; DB 4; Length 17;
Best Local Similarity 80.0%; Pred. No. 14;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 2 AGDIK 6
   |||:|
Db 13 AGDVK 17
```

;; TITLE OF INVENTION: Acid Dioxygenase
;; NUMBER OF SEQUENCES: 2
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Quarles & Brady
;; STREET: PO Box 2113
;; CITY: Madison
;; STATE: WI
;; COUNTRY: USA
;; ZIP: 53701-2113
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/592,900
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Seay, Nicholas J
;; REGISTRATION NUMBER: 27,386
;; REFERENCE/DOCKET NUMBER: 920214.90158
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 608-251-5000
;; TELEFAX: 608-251-9166
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 445 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-592-900-2

Query Match 86.28; Score 25; DB 3; Length 445;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6
Db 108 SAGEIK 113

RESULT 8
US-08-979-917A-1
; Sequence 1, Application US/08979917A
; Patent No. 6118050
; GENERAL INFORMATION:
; APPLICANT: STURNER, STEPHEN
; APPLICANT: HIRAYAMA, LYNNE MIYO
; APPLICANT: SINGH, BIJAY
; APPLICANT: BASCOMB, NEWELL
; TITLE OF INVENTION: HPD GENE AND INHIBITORS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,917A
; FILING DATE: 25-JUL-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/022,604
; FILING DATE: 25-JUL-1996

;; ATTORNEY/AGENT INFORMATION:
;; NAME: Zitron, Anne E.
;; REGISTRATION NUMBER: 41,391
;; REFERENCE/DOCKET NUMBER: 0646/1B917-US1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-527-7700
;; TELEFAX: 212-753-6237
;; TELFX: 236687
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 446 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: No. 6118050e
;; US-08-979-917A-1

Query Match 86.2%; Score 25; DB 3; Length 446;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6
Db 108 SAGEIK 113

RESULT 9
US-08-378-761A-67
; Sequence 67, Application US/08378761A
; Patent No. 5635384
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,761A
; FILING DATE: 26-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-378-761A-67

Query Match 82.8%; Score 24; DB 1; Length 11;
Best Local Similarity 80.0%; Pred. No. 9;

OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,563
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/030,601
FILING DATE: NOVEMBER 13, 1996
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9982
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 18:
LENGTH: 709 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORGANISM: GUT1
US-08-968-563-18

Query Match 89.7%; Score 26; DB 3; Length 709;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
Db 105 NAGDIK 110

RESULT 5
US-08-644-664B-14
Sequence 14, Application US/08644664B
Patent No. 5776746
GENERAL INFORMATION:
APPLICANT: Denney Jr., Dan W.
TITLE OF INVENTION: Gene Amplification Methods
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/644,664B
FILING DATE: 01-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: GENITOPE-00912
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 218 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-644-664B-14
Query Match 86.2%; Score 25; DB 1; Length 218;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 SAGDIK 6
Db 110 STGDIK 115
RESULT 6
US-08-761-277A-14
Sequence 14, Application US/08761277A
Patent No. 5972334
GENERAL INFORMATION:
APPLICANT: Denney Jr., Dan W.
TITLE OF INVENTION: Vaccines For Treatment Of Lymphoma And
TITLE OF INVENTION: Leukemia
NUMBER OF SEQUENCES: 80
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/761,277A
FILING DATE: 06-DEC-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/644,664
FILING DATE: 01-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: MacKnight, Kamrin T.
REGISTRATION NUMBER: 38,230
REFERENCE/DOCKET NUMBER: GENITOPE-02406
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 218 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-761-277A-14

Query Match 86.2%; Score 25; DB 2; Length 218;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SAGDIK 6
Db 110 STGDIK 115

RESULT 7
US-08-592-900-2
Sequence 2, Application US/08592900
Patent No. 6087563
GENERAL INFORMATION:
APPLICANT: Dellapenna, Dean
TITLE OF INVENTION: Cloned Plant P-Hydroxyphenyl Pyruvic

Db 37 SAGDIK 42

RESULT 2

US-08-468-011A-2
; Sequence 2, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Steven
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1744
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-011A-2

Query Match 100.0%; Score 29; DB 3; Length 541;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6

|||||

Db 310 SAGDIK 315

RESULT 3

PCT-US95-07085-2
; Sequence 2, Application PC/TUS9507085
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Steven
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone Receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland

; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07085
; FILING DATE: 05-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-393
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-07085-2

Query Match 100.0%; Score 29; DB 4; Length 541;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6

|||||

Db 310 SAGDIK 315

RESULT 4

US-08-968-563-18
; Sequence 18, Application US/08968563
; Patent No. 6013494
; GENERAL INFORMATION:
; APPLICANT: CHARLES E. NAKAMURA
; APPLICANT: ANTHONY A. GATENBY
; APPLICANT: AMY (KUANG-HUA) HSU
; APPLICANT: RICHARD D. LA REAU
; APPLICANT: SHARON L. HAYNIE
; APPLICANT: MARIA DIAZ-TORRES
; APPLICANT: DONALD E. TRIMBUR
; APPLICANT: GREGORY M. WHITED
; APPLICANT: VASANTHA NAGARAJAN
; APPLICANT: MARK S. PAYNE
; APPLICANT: STEPHEN K. PICATAGGIO
; APPLICANT: RAMESCH V. NAIR
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
; TITLE OF INVENTION: PRODUCTION OF 1,3-PROPANEDIOL
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: U.S.A.
; ZIP: 19898
; ADDRESSEE: GENENCOR INTERNATIONAL, INC.
; STREET: 4 CAMBRIDGE PLACE
; STREET: 1870 SOUTH WINTON ROAD
; CITY: ROCHESTER
; STATE: NEW YORK
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.50 INCH DISKETTE
; COMPUTER: IBM PC COMPATIBLE

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:49:18 ; Search time 97.15 Seconds
(without alignments)
1.035 Million cell updates/sec

Title: US-09-236-468A-2_COPY_310_315

Perfect score: 29

Sequence: 1 SAGDIK 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 164575

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	60	3	US-08-468-011A-11
2	29	100.0	541	3	US-08-468-011A-2
3	29	100.0	541	4	PCT-US95-07085-2
4	26	89.7	709	3	US-08-968-563-18
5	25	86.2	218	1	US-08-644-664B-14
6	25	86.2	218	2	US-08-761-277A-14
7	25	86.2	445	3	US-08-592-900-2
8	25	86.2	445	3	US-08-979-917A-1
9	24	82.8	11	1	US-08-378-761A-67
10	24	82.8	11	1	US-08-485-286-67
11	24	82.8	17	4	PCT-US93-08739-2
12	24	82.8	19	2	US-08-374-560-4
13	24	82.8	30	1	US-08-446-692-86
14	24	82.8	30	2	US-08-488-351A-86
15	24	82.8	42	4	PCT-US93-08739-3
16	24	82.8	48	3	US-08-938-830-3
17	24	82.8	48	3	US-09-020-222-3
18	24	82.8	61	2	US-08-374-560-1
19	24	82.8	65	1	US-08-446-692-111
20	24	82.8	65	2	US-08-488-351A-111
21	24	82.8	250	1	US-07-854-845B-8
22	24	82.8	250	1	US-07-901-707-8
23	24	82.8	250	1	US-07-988-430-8
24	24	82.8	250	1	US-08-425-336-8
25	24	82.8	250	1	US-08-378-761A-76
26	24	82.8	250	1	US-08-485-286-76
27	24	82.8	250	1	US-08-488-113B-8
28	24	82.8	250	1	US-08-477-484B-8

29 24 82.8 250 2 US-08-646-360-8 Sequence 8, Appli
30 24 82.8 250 4 PCT-US92-09487-8 Sequence 8, Appli
31 24 82.8 396 2 US-08-838-219B-9 Sequence 9, Appli
32 24 82.8 396 3 US-09-233-336A-9 Sequence 9, Appli
33 24 82.8 400 3 US-08-938-830-29 Sequence 29, Appli
34 24 82.8 415 3 US-08-938-830-1 Sequence 1, Appli
35 24 82.8 415 3 US-09-020-222-1 Sequence 1, Appli
36 24 82.8 488 1 US-08-243-542-1 Sequence 1, Appli
37 24 82.8 488 1 US-08-477-407-1 Sequence 1, Appli
38 24 82.8 488 1 US-08-484-355-1 Sequence 1, Appli
39 24 82.8 524 1 US-08-243-542-2 Sequence 2, Appli
40 24 82.8 524 1 US-08-477-407-2 Sequence 2, Appli
41 24 82.8 524 1 US-08-484-355-2 Sequence 2, Appli
42 24 82.8 626 1 US-08-472-934-6 Sequence 6, Appli
43 24 82.8 626 2 US-08-323-460A-6 Sequence 6, Appli
44 24 82.8 626 2 US-08-461-146C-6 Sequence 6, Appli
45 24 82.8 626 3 US-08-461-145C-6 Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-468-011A-11
; Sequence 11, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-011A-11

Query Match 100.0%; Score 29; DB 3; Length 60;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SAGDIK 6
|||||

SQ Sequence 218 AA;

Query Match 86.2%; Score 25; DB 18; Length 218;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SAGDIK 6
|
Db 110 stgdik 115

RESULT 15

R05026
ID R05026 standard; protein; 243 AA.

XX AC R05026;

XX DT 02-OCT-1990 (first entry)

XX DE Beta subunit of rat high affinity IgE receptor Fc(epsilon)RI.

XX KW high affinity IgE receptor; beta subunit of IgE receptor; rat;
XX KW allergic response; ss.

XX OS synthetic.

XX FH Key Location/Qualifiers

XX FT Domain 60...79

XX FT Domain /label-putative transmembrane domain

XX FT Domain 98...117

XX FT Domain /label-putative transmembrane domain

XX FT Domain 131...150

XX FT Domain /label-putative transmembrane domain

XX FT Domain 180...199

XX FT Domain /label-putative transmembrane domain

XX PN W09004640-A.

XX PD 03-MAY-1990.

XX PF 18-OCT-1989; 89WO-US04628.

XX PR 18-OCT-1988; 88US-0259065.

XX PA (USDC) US SEC OF COMMERCE.

XX PI Kinet JP, Metzger H;

XX DR WPI; 1990-164023/21.

XX DR N-PSDB; Q04645.

XX PT DNA sequences for subunit peptides of high affinity IgE receptor -
XX PT and derived polypeptides, for therapy and diagnosis of
XX PT allergies, and studies of IgE receptor interaction

XX PS Disclosure; ; Opp; English.

XX CC The high affinity receptor is a tetrameric complex consisting of 2
XX CC gamma subunits and one each of subunits alpha and beta. It is
XX CC expressed on mast cells and is involved in the allergic response.
XX CC COS-7 cells cotransfected with cDNA for all 3 subunit types
XX CC (derived from rat basophilic leukaemia cells) express receptor on
XX CC their surfaces. Detailed study of the receptors is now possible.
XX CC See also Q04643-4 and Q04645.

XX SQ Sequence 243 AA;

Query Match

86.2%; Score 25; DB 11; Length 243;

Best Local Similarity 83.3%; Pred. No. 2.2e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

XX Homo sapiens.

XX WO9902546-A1.

XX 21-JAN-1999.

XX 07-JUL-1998; 98WO-US13684.

XX 12-SEP-1997; 97US-0058785.

XX 08-JUL-1997; 97US-0051916.

XX 08-JUL-1997; 97US-0051918.

XX 08-JUL-1997; 97US-0051919.

XX 08-JUL-1997; 97US-0051920.

XX 08-JUL-1997; 97US-0051925.

XX 08-JUL-1997; 97US-0051926.

XX 08-JUL-1997; 97US-0051928.

XX 08-JUL-1997; 97US-0051929.

XX 08-JUL-1997; 97US-0051930.

XX 08-JUL-1997; 97US-0051931.

XX 08-JUL-1997; 97US-0051932.

XX 08-JUL-1997; 97US-0052732.

XX 08-JUL-1997; 97US-0052733.

XX 08-JUL-1997; 97US-0052793.

XX 08-JUL-1997; 97US-0052795.

XX 08-JUL-1997; 97US-0052803.

XX 18-AUG-1997; 97US-0055684.

XX 18-AUG-1997; 97US-0055722.

XX 18-AUG-1997; 97US-0055723.

XX 18-AUG-1997; 97US-0055947.

XX 18-AUG-1997; 97US-0055948.

XX 18-AUG-1997; 97US-0055949.

XX 18-AUG-1997; 97US-0055950.

XX 18-AUG-1997; 97US-0055953.

XX 18-AUG-1997; 97US-0055954.

XX 18-AUG-1997; 97US-0055964.

XX 18-AUG-1997; 97US-0055984.

XX 18-AUG-1997; 97US-0056360.

XX 12-SEP-1997; 97US-0058660.

XX 12-SEP-1997; 97US-0058661.

XX 12-SEP-1997; 97US-0058664.

(HUMA-) HUMAN GENOME SCI INC.

Brewer LA, Ebner R, Fischer CL, Kyaw H, Lafleur DW, Li Y, Moore PA;

Olsen HS, Rosen CA, Ruben SM, Shi Y, Soppet DR, zeng Z;

WPI; 1999-120770/10.

New isolated human genes and the secreted polypeptides they encode -
useful for diagnosis and treatment of e.g. cancers, neurological
disorders, immune diseases, inflammation or blood disorders

Disclosure; Page 61; 464pp; English.

This sequence represents a fragment of a secreted human protein encoded
by the nucleic acid molecule detailed in the descriptor line. The gene
can be used to generate fusion proteins by linking to the gene to a human
immunoglobulin Fc portion (e.g. X27302) for increasing the stability of
the fused protein as compared to the human protein only.

The invention relates to 123 novel genes and their fragments (nucleic
acid sequences: X27311-X27449; amino acid sequences Y02650-Y02786) which
are useful for preventing, treating or ameliorating medical conditions
e.g. by protein or gene therapy. Also, pathological conditions can be

CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 123 polynucleotides, based on
CC which tissues they are most highly expressed in (see X27311 for described
CC uses).

XX Sequence 187 AA;

Query Match 86.2%; Score 25; DB 20; Length 187;

Best local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 AGDIK 6

Db 79 agdik 83

RESULT 14

W37339

ID W37339 standard; Protein; 218 AA.

XX W37339;

XX 11-MAY-1998 (first entry)

XX Mouse HPRT.

XX Vaccine; B-cell malignancy; lymphoma; leukaemia; tumour;
KW gene amplification; immunotherapy; therapy; mouse; HPRT;
KW hypoxanthine-phosphoribosyltransferase; selectable marker;
KW vector; pMSD5-HPRT.

XX Mus musculus.

XX W09741244-A1.

XX 06-NOV-1997.

XX 25-APR-1997; 97WO-US07039.

XX 06-DEC-1996; 96US-0761277.

XX 01-MAY-1996; 96US-0644664.

XX (GENI-) GENITOPE CORP.

XX Denney DW;

XX WPI; 1997-549743/50.

XX N-PSDB; T97166.

XX Multivalent vaccine to treat B cell lymphoma or leukaemia -
comprises at least 2 different recombinant variable regions of
immunoglobulin molecules derived from B cell lymphoma cells

Example 2; Page 108-109; 177pp; English.

This sequence comprises hypoxanthine phosphoribosyltransferase
(HPRT) from mouse. The HPRT gene (see T97166) was utilised in
the construction of selection vector pMSD5-HPRT, which contains
full-length HPRT DNA under transcriptional control of the Moloney
LTR (see also T97160). The invention provides vectors and improved
methods for the expression and co-amplification of genes encoding
recombinant proteins in cultured cells. The methods permit the
isolation of cell lines which have co-amplified input recombinant
sequences which encode an amplifiable marker, one or more
expression vectors encoding a protein of interest and optionally a
selectable marker. The amplified cells provide large quantities of
recombinant proteins suitable for immunotherapy for treatment of
lymphomas and leukaemias. The methods permit the production of
custom vaccines, including multivalent vaccines, that reflect the
degree of somatic variation found in a patient's tumour.

Query Match 86.2%; Score 25; DB 20; Length 67;
 Best Local Similarity 83.3%; Pred. No. 54;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SAGDIK 6
 I I I I I
 Db 7 stgdik 12

RESULT 11
 R05027
 ID R05027 standard; protein; 112 AA.
 XX
 AC R05027;
 XX
 DT 02-OCT-1990 (first entry)
 XX
 DE Deleted form of beta subunit of rat high affinity IgE receptor.
 XX
 KW high affinity IgE receptor; deleted form of beta subunit of IgE receptor;
 KW rat; allergic response; ss.
 XX
 OS synthetic.
 XX
 PN W09004640-A.
 XX
 PD 03-MAY-1990.
 XX
 PF 18-OCT-1989; 89WO-US04628.
 XX
 PR 18-OCT-1988; 88US-0259065.
 XX
 PA (USDC) US SEC OF COMMERCE.
 XX
 PI Kinet JP, Metzger H;
 XX
 DR WPI; 1990-164023/21.
 DR N-PSDB; Q04646.

XX DNA sequences for subunit peptides of high affinity IgE receptor -
 PT and derived polypeptides, for therapy and diagnosis of
 PT allergies, and studies of IgE receptor interaction
 XX
 PS Disclosure; : Opp; English.
 XX
 CC This sequence is encoded by a deleted form of the cDNA encoding the beta
 CC subunit. The protein is consequently a truncated form of the intact
 CC protein. (see R05026).
 CC The high affinity receptor is a tetrameric complex consisting of 2
 CC gamma subunits and one each of subunits alpha and beta. It is
 CC expressed on mast cells and is involved in the allergic response.
 CC COS-7 cells cotransfected with cDNA for all 3 intact subunit types
 CC (derived from rat basophilic leukaemia cells) express receptor on
 CC their surfaces. Detailed study of the receptors is now possible.
 CC See also Q04643-5.
 XX
 SQ Sequence 112 AA;

Query Match 86.2%; Score 25; DB 11; Length 112;
 Best Local Similarity 83.3%; Pred. No. 95;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
 I I I I I
 Db 21 sagdie 26

RESULT 12
 R14771
 ID R14771 standard; Protein; 115 AA.
 XX
 AC R14771;
 DE

XX 03-FEB-1992 (first entry)
 DT
 XX Truncated beta subunit of high affinity IgE receptor.
 DE
 XX Immunoglobulin; receptor; Fc(epsilon)RI.
 KW
 XX Rattus.
 OS

XX Key Location/Qualifiers
 XX Region 63..82
 FT /label= transmembrane
 FT /note= "putative"
 FT Peptide 14..32
 FT /label= tryptic
 FT /note= "sequenced directly"
 FT Peptide 40..46
 FT /label= tryptic
 FT /note= "sequenced directly"

XX US7626704-A.
 PN
 XX 15-OCT-1991.
 PD
 XX 14-DEC-1990; 90US-0151091.
 PF
 XX 14-DEC-1990; 90US-0626704.
 PR
 XX (USSH) NAT INST OF HEALTH.
 PA
 XX Kinet JP, Metzger H;
 PI
 XX WPI; 1991-346755/47.
 DR N-PSDB; Q14735.

XX DNA coding alpha, beta and gamma-units of IgE high affinity
 PT receptor - are used to prepare recombinant polypeptide(s) for
 PT treating allergy, drug screening or monitoring IgE level
 XX
 PS Disclosure; Fig 6B; 58pp; English.
 XX

XX A lambda gtl library was prepared from poly-A RNA isolated from rat
 CC basophilic leukaemia cells. Probes were designed based on the
 CC sequence of a tryptic peptide of Fc(epsilon)RI beta-subunit. The
 CC longest clone was sequenced (see Q14734). It predicts two possible
 CC initiation codons. To check that the true start codon was not still
 CC further 5', the cDNA library was rescreened. Of the additional 28
 CC clones which were isolated, 6 showed early termination; this
 CC sequence was deduced from the truncated coding sequence. The first 3
 CC amino acid residues may be absent, depending upon which of the two
 CC initiation codons is actually used.

XX Sequence 115 AA;

Query Match 86.2%; Score 25; DB 12; Length 115;
 Best Local Similarity 83.3%; Pred. No. 98;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SAGDIK 6
 I I I I I
 Db 24 sagdie 29

RESULT 13
 Y02844
 ID Y02844 standard; Protein; 187 AA.
 XX
 AC Y02844;

XX
 DT 11-JUN-1999 (first entry)
 XX
 DE Fragment of human secreted protein encoded by gene 48.

XX PS Disclosure; Fig 4A-C; 52pp; English.

XX CC The present sequence represents avian topoisomerase I. The

XX CC invention provides vectors containing the topoisomerase DNA. The

XX CC host cells that express the topoisomerase I at high levels, and

XX CC methods for the recombinant production of topoisomerase I. The

XX CC invention provides a claimed method of improving the efficiency

XX CC of a reverse transcriptase (RT) based molecular reaction by

XX CC including topoisomerase I in a reaction mixture. The molecular

XX CC reaction is preferably a RT-based cDNA synthesis, the RT is derived

XX CC from lentivirus and the topoisomerase I is human, murine or avian

XX CC (see Y06424-26). cDNA synthesis kits are claimed. Topoisomerase

XX CC I can also be used to improve the integration efficiency of

XX CC retrovirus vectors and hence to improve retrovirus-based gene

XX CC therapy and gene delivery techniques, and also to improve the

XX CC efficiency of RT-mediated PCR.

XX SQ Sequence 766 AA;

Query Match 89.7%; Score 26; DB 20; Length 766;

Best Local Similarity 83.3%; Pred. No. 4.4e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6

Db 100 ssgdik 105

RESULT 9

Y52991

ID Y52991 standard; Protein; 924 AA.

XX AC Y52991;

XX 21-FEB-2000 (first entry)

XX Drosophila melanogaster trachealess (trh) protein.

XX Drosophila melanogaster; trachealess; trh; gene screening; cell death;

KW apoptosis; cancer; immune system; mutation.

XX OS Drosophila melanogaster.

XX WO99555906-A2.

XX 04-NOV-1999.

XX 26-APR-1999; 99WO-CA00377.

XX 27-APR-1998; 98US-0083077.

XX (TORO-) TORONTO HOSPITAL.

XX Woodgett JR, Manoukian AS, Jing J, Wetsch B;

XX WPI; 2000-052812/04.

XX N-PSDB; 232295.

XX Determining the presence of a mutation capable of enhancing programmed

XX cell death in Drosophila melanogaster.

XX Claim 79; Page 57; 58pp; English.

XX A method has been developed for determining whether a mutation capable

XX of enhancing programmed cell death (PCD) is present in a Drosophila

XX melanogaster fly; The method comprises: (1) crossing a first mutant fly,

XX having a recessive first mutation, which is known to induce PCD in an

XX embryo of a fly homozygous for the mutation, with a second mutant fly

XX for determination; and (2) determining whether the embryos obtained have

XX undergone ectopic PCD. The method can be used to determine whether a

XX mutation is capable of disrupting development in D. melanogaster and

CC also whether it is present in a particular fly. PCD can be induced by

CC administration of antisense nucleic acids sufficiently complementary to

CC inhibit translation of DAKt1 or PP2A mRNA present in a cell.

CC D. melanogaster embryos can be used to determine whether a substance is

CC useful for inducing PCD or inhibiting cell death. Determining the

CC presence of a nucleotide sequence encoding a trachealess protein, DAKt1

CC or PP2A capable of initiating transcription is used to determine a

CC predisposition for conferring abnormal development on offspring of the

CC subject containing the sequence. An understanding of PCD is important

CC for determining possible new therapeutics useful against, e.g. cancer

CC or an overactive immune system. The present sequence represents the

CC trachealess protein which is capable of initiating transcription, and

CC can be used in the method from the present invention.

XX SQ Sequence 924 AA;

Query Match 89.7%; Score 26; DB 21; Length 924;

Best Local Similarity 83.3%; Pred. No. 5.4e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6

Db 521 sagdmk 526

RESULT 10

Y42644

ID Y42644 standard; Protein; 67 AA.

XX AC Y42644;

XX 10-JAN-2000 (first entry)

XX Brassica napus DZ2B partial fragment.

XX Signal transduction protein; dehiscence; male sterile plant; DZ2B gene;

KW shatter resistance; oilseed rape; response regulator protein.

XX OS Brassica napus.

XX WO9949046-A1.

XX 30-SEP-1999.

XX 22-MAR-1999; 99WO-GB009005.

XX 20-MAR-1998; 98GB-0006113.

XX (BIOG-) BIOGEMMA UK LTD.

XX Wyatt P, Roberts JA, Whitelaw C;

XX WPI; 1999-580449/49.

XX N-PSDB; 222975.

XX A nucleic acid encoding a signal transduction protein involved in plant

XX dehiscence, useful for producing shatter resistant male sterile plants

XX Example 2; Fig 5; 71pp; English.

XX The invention provides a nucleic acid encoding a signal transduction

XX protein involved in the process of dehiscence. The nucleic acids and

XX proteins are useful for regulating or controlling dehiscence of a pod or

XX an anther in a plant, useful in the production of male sterile plants.

XX The methods, etc. may be used in the production of shatter resistance or

XX shatter-delayed plants such as oilseed rape (Brassica napus). The present

XX sequence represents the partial fragment of B. napus DZ2B protein.

XX SQ Sequence 67 AA;

XX 10-NOV-1997; 97WO-US20293.
 XX 13-NOV-1996; 96US-0030602.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA (GEMV) GENENCOR INT INC.
 XX Bulthuis BA, Gatenby AA, Haynie SL, Hsu AK, Lareau RD;
 PI WPI: 1998-297943/26.
 DR N-PSDB; V30827.
 XX Fermentative production of glycerol using recombinant host -
 PT containing genes for glycerol-3-phosphate dehydrogenase and/or
 PT glycerol-3- phosphate
 XX Claim 11; Page 43-46; 57pp; English.
 XX This Saccharomyces polypeptide comprises a cytosolic
 CC glycerol kinase that catalyses the conversion of glycerol to
 CC glycerol-3-phosphate. It is encoded by the GUT1 gene (see V30827).
 CC The invention provides recombinant organisms that express glycerol
 CC 3-phosphatase (G3P) and/or glycerol-3-phosphate dehydrogenase
 CC (G3PDH) (see also W57324-32) useful for the production of glycerol
 CC from a variety of C-sources. A host cell is preferably transformed
 CC with a cassette containing either a G3PDH gene and/or a G3P gene
 CC and then cultured in the presence of a mono-, oligo-, polysaccharide
 CC or 1C-substrate. The glycerol obtained is used in cosmetics, liquid
 CC soaps, pharmaceuticals, lubricants and antifreezes; its esters are
 CC used in the oil and fat industries. The method produces glycerol
 CC rapidly and inexpensively without generation of polluting
 CC by-products.
 XX SQ Sequence 709 AA;
 Query Match 89.7%; Score 26; DB 19; Length 709;
 Best Local Similarity 83.3%; Pred. No. 4.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SAGDIK 6
 Db 105 nagdik 110
 RESULT 7
 ID Y26171 standard; Protein; 709 AA.
 XX AC Y26171;
 XX 29-SEP-1999 (first entry)
 XX Cytosolic glycerol kinase encoded by GUT1 gene.
 DE GUT1; glycerol kinase; glycerol; recombinant organism; transformation;
 KW glycerol biosynthetic pathway; expression cassette; 1-3 propanediol;
 KW pharmaceutical compound; antifreeze solution; lubricant; polyurethane;
 KW cyclic compound; fat and oil industry; polyester fiber;
 KW glycerol-3-phosphate dehydrogenase; glycerol-3-phosphatase.
 XX OS Saccharomyces sp.
 XX WO9928480-A1.
 XX 10-JUN-1999.
 XX 02-DEC-1998; 98WO-US25551.
 XX 02-DEC-1997; 97US-0982783.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.

PA (GEMV) GENENCOR INT INC.
 XX Nair RV, Payne MS, Trimbur DE, Valle F;
 XX WPI: 1999-385384/32.
 DR N-PSDB; X80616.
 XX Recombinant organisms containing G3PDH and or G3P phosphatase
 PT Disclosure; Pages 70-73; 84pp; English.
 XX The present sequence is a cytosolic glycerol kinase encoded by GUT1
 CC gene which catalyses the conversion of glycerol and ATP to glycerol-3
 CC -phosphate and ADP. This is used in the production of glycerol from a
 CC recombinant organism by transforming a suitable host cell with an
 CC expression cassette comprising either one or both of the genes encoding
 CC G3PDH and G3P, where the host cell has disruptions in either glycerol
 CC kinase or glycerol dehydrogenase endogenous genes to prevent their
 CC active expression. The transformed host cell is cultured with a carbon
 CC source and glycerol is recovered. Compounds derived from the glycerol
 CC biosynthetic pathway like 1,3-propanediol can also be produced. The
 CC method provides a rapid, inexpensive and environment-friendly source of
 CC glycerol. Glycerol is used in cosmetics, food, pharmaceuticals,
 CC lubricants, anti-freeze solutions, fat and oil industry etc.. 1,3
 CC -propanediol is used for the production of polyester fibers and the
 CC manufacture of polyurethanes and cyclic compounds.
 XX SQ Sequence 709 AA;
 Query Match 89.7%; Score 26; DB 20; Length 709;
 Best Local Similarity 83.3%; Pred. No. 4.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SAGDIK 6
 Db 105 nagdik 110
 RESULT 8
 ID Y06426 standard; Protein; 766 AA.
 XX AC Y06426;
 XX 27-SEP-1999 (first entry)
 XX Avian topoisomerase I.
 DE Topoisomerase I; bird; reverse transcriptase; retrovirus; vector;
 KW gene therapy; gene delivery; PCR.
 XX Aves.
 XX WO9935254-A1.
 XX 15-JUL-1999.
 XX 08-JAN-1999; 99WO-US00559.
 XX 20-FEB-1998; 98US-0075622.
 XX 08-JAN-1998; 98US-0072293.
 XX (HALL/) HALL W W.
 PA (SONG/) SONG E S.
 XX Hall WW, Song ES;
 XX WPI: 1999-430390/36.
 DR N-PSDB; X59397.
 XX Use of topoisomerase I for improving efficiency of reverse
 PT transcriptase based molecular reactions

Db 377 sagdlk 382

RESULT 4

W30685
ID W30685 standard; Protein; 709 AA.

AC W30685;

XX 12-OCT-1998 (first entry)

XX Glycerol kinase GUT1.

XX Glycerol kinase; gut1 gene; 1,3-propanediol.
XX Saccharomyces sp.

OS W09821339-A1.

XX 22-MAY-1998.

XX 10-NOV-1997; 97WO-US20292.

XX 13-NOV-1996; 96US-0030601.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX (GEMV) GENENCOR INT INC.

XX Dias-Torres M, Galenby AA, Haynie SL, Hsu AK, Lareau RD;

XX Nagarajan V, Nair RV, Nakamura CE, Payne MS, Picataggio SK;

XX Trimbur DB, Whited GM;

XX WPI; 1998-297942/26.

XX N-PSDB; V42021.

XX Fermentative production of 1,3-propanediol - by single organism
XX containing cassette comprising specific genes, and capable of using
XX inexpensive carbon sources
XX Claim 11; Page 65-67; 95pp; English.
XX Cytosolic glycerol kinase GUT1 is encoded by the gut1 gene (see
XX (V42021) of Saccharomyces sp. This enzyme catalyses the conversion
XX of glycerol to glycerol-3-phosphate, or the reverse reaction. A
XX claimed method for production of 1,3-propanediol (I) comprises
XX culturing a microorganism transformed with a cassette containing at
XX least 1 of the genes (see V42012-21) for glycerol-3-phosphate
XX dehydrogenase, glycerol-3-phosphatase, glycerol dehydratase and
XX 1,3-propanediol oxidoreductase (see W30676-85). Any of these genes
XX not present on the cassette must be present endogenously. The
XX enzyme sequences may include substitutions, deletions and additions
XX provided activity is not altered. A single recombinant organism
XX can now be used for production of (I) from inexpensive C-sources
XX (contrast use of glycerol or dihydroxyacetone) without causing
XX pollution. (I) is a starting material for polyesters, polyurethanes
XX and cyclic compounds.

XX Sequence 709 AA;

Query Match

Best Local Similarity 89.7%; Score 26; DB 19; Length 709;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6

Db 105 nagdlk 110

RESULT 5

W60262
ID W60262 standard; Protein; 709 AA.

XX

AC W60262;

XX 28-SEP-1998 (first entry)

XX Klebsiella pneumoniae glycerol kinase.

XX glycerol kinase; production; 1,3-propanediol; recombinant.

XX Klebsiella pneumoniae.

XX W09821341-A2.

XX 22-MAY-1998.

XX 13-NOV-1997; 97WO-US20873.

XX 13-NOV-1996; 96US-0030601.

XX (GEMV) GENENCOR INT INC.

XX Chase MW, Diaz-torres M, Dunn-coleman NS, Trimbur D;

XX WPI; 1998-297944/26.

XX N-PSDB; V35742.

XX New method for increasing production of 1,3-propanediol - comprises
XX fermentation of inexpensive carbon sources by microorganism
XX expressing dehydratase, used, e.g. to prolong half-life of enzyme
XX Disclosure; Page 77-79; 133pp; English.
XX The sequence is that of cytosolic glycerol kinase.
XX It was used as part of a method of fermentative production
XX of 1,3-propanediol (1,3-pd), using an organism comprising
XX at least 1 gene encoding a dehydratase, is improved by
XX inserting into the host a gene encoding protein X and culturing
XX the transformant in presence of a carbon source (e.g. mono-, oligo-
XX or poly-saccharide or LC substrate) convertible to 1,3-pd.
XX 1,3-pd is a starting material for polyesters, polyurethanes and
XX cyclic compounds. 1,3-pd can now be produced by a single
XX recombinant organism from inexpensive carbon sources such
XX as glucose (rather than costly glycerol or dihydroxyacetone),
XX rapidly and without causing pollution.

XX Sequence 709 AA;

Query Match

Best Local Similarity 89.7%; Score 26; DB 19; Length 709;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6

Db 105 nagdlk 110

RESULT 6

W57329
ID W57329 standard; Protein; 709 AA.

XX W57329;

XX 14-SEP-1998 (first entry)

XX Mitochondrial glycerol kinase GUT1.

XX Glycerol kinase; GUT1; yeast.

XX Saccharomyces sp.

XX W09821340-A1.

XX 22-MAY-1998.

PT hyper-calcaemia, hypo- or hyper-phosphatemia, kidney stones, etc

PS Claim 9; Fig 1A-E; 62pp; English.

CC A novel 7-transmembrane receptor (W12695) has been identified as a

CC human G-protein parathyroid hormone (PTH) receptor, designated

CC H12DG74. It shows 48.2% homology to the human PTH receptor. Its

CC amino acid sequence was deduced from a cDNA clone (159619) isolated

CC from a human T cell lymphoma tissue cDNA library. Recombinant

CC H12DG74 can be produced in transformed host cells and used to

CC screen for (ant)agonist cpds. Agonists may be used to prevent or

CC treat e.g. hypocalcaemia, hyperphosphataemia, hypoparathyroidism

CC and chronic tetany by stimulating an increase in serum calcium

CC levels. Antagonists can be used to inhibit the receptor e.g. for

CC the treatment of osteoporosis, hypercalcaemia, hypoparathyroidism,

CC hypophosphataemia, kidney stone, nephrolithiasis.

XX

SQ Sequence 541 AA;

Query Match 100.0%; Score 29; DB 18; Length 541;

Best Local Similarity 100.0%; Pred. No. 61;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6

Db 310 sagdik 315

RESULT 2

W56322

ID W56322 standard; Protein; 2039 AA.

AC W56322;

XX

DT 19-AUG-1998 (first entry)

XX

DE Haemophilus paragallinarum antigenic protein #2.

XX

DE Haemophilus paragallinarum; antigenic protein; HI antibody; diagnosis;

KW vaccine; chicken infectious coryza; CIC; fowl.

KW

XX

OS Haemophilus paragallinarum.

XX

FH Key Location/Qualifiers

FT Peptide 1..70

FT /label= signal

FT Protein 71..2039

FT /note= "antigenic protein"

FT

XX W09812331-A1.

PN

XX

PD 26-MAR-1998.

XX

PF 12-SEP-1997; 97WO-JP03222.

XX

PR 19-SEP-1996; 96JP-0271408.

XX

PA (KAGA) ZH KAGAKU & KESSEI RYOHO KENKYUSHO.

PA (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.

XX

PI Hamada F, Matsuo K, Sakaguchi M, Tokiyoshi S, Tokunaga E;

XX

DR WPI: 1998-230318/20.

DR N-PSDB: V22837.

XX

XX Antigenic polypeptide from Haemophilus paragallinarum induces HI

PT antibody production - and is useful for diagnosis of and preparation

PT of vaccines for chicken infectious coryza

XX

PS Claim 5; Page 71-87; 108pp; Japanese.

XX

CC The present sequence represents an antigenic protein derived from

CC Haemophilus paragallinarum strain C-53-47. The antigenic protein

CC stimulates the production of HI antibodies in fowl. The protein

CC and DNA coding for it can be used in the preparation of vaccines

CC for the prevention of chicken infectious coryza (CIC). The protein

CC and its antibodies can be used in the diagnosis and treatment of CIC.

XX

SQ Sequence 2039 AA;

Query Match 100.0%; Score 29; DB 19; Length 2039;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6

Db 489 sagdik 494

RESULT 3

R14163

ID R14163 standard; Protein; 445 AA.

XX

AC R14163;

XX

DT 09-DEC-1991 (first entry)

XX

DE Cellular DNA-binding protein R51.

XX

KW Epidermal keratinocyte; tandem repeat; E2; subregion; HPV 18;

KW human papilloma virus; cancer; URR; upstream regulatory region.

XX

OS Homo sapiens.

XX

PN EP449170-A.

XX

PD 02-OCT-1991.

XX

PF 25-MAR-1991; 91EP-0104659.

XX

PR 30-MAR-1990; 90DE-4010237.

XX

PA (BEHW) BEHRINGERWERKE AG.

XX

PI Hoppe-Seyle F, Hirt L, Butz K, Bartsch D, Bauknecht T;

PI Royer H;

XX

DR WPI: 1991-289624/40.

DR N-PSDB: Q13880.

XX

PT Cellular deoxyribonucleic acid binding protein R51 - useful for

PT treating cancer caused by human papilloma virus, and DNA coding

PT for it

XX

PS Disclosure; Fig 3(A-B); 29pp; German.

XX

CC DNA encoding R51 has been isolated from a lambda gt11 expression

CC library contg. cDNA from primary human epidermal keratinocytes by

CC screening with an oligonucleotide probe comprising two tandem repeats

CC of the viral E2 transactivator/transrepressor recognition region.

CC The R51 protein binds specifically to the 28 bp E2 subregion of

CC the upstream regulatory region (URR) of the human papilloma virus

CC HPV 18. They may be useful in the treatment of HPV-induced cancer.

CC See also Q13881-89.

XX

SQ Sequence 445 AA;

Query Match 93.1%; Score 27; DB 12; Length 445;

Best Local Similarity 83.3%; Pred. No. 1.4e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SAGDIK 6

Db 11111

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:51:43 ; Search time 138.73 Seconds
(without alignments)
1.479 Million cell updates/sec

Title: US-09-236-468A-2_COPY_310_315
Perfect score: 29
Sequence: 1 SAGDIK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36: *
1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT: *
2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT: *
3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT: *
4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT: *
5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT: *
6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT: *
7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT: *
8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT: *
9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT: *
10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT: *
11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT: *
12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT: *
13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT: *
14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT: *
15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT: *
16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT: *
17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT: *
18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT: *
19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT: *
20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT: *
21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	100.0	541	18 W12695	G-protein parathyr
2	29	100.0	2039	19 W56322	Haemophilus paraga
3	27	93.1	445	12 R14163	Cellular DNA-bind
4	26	89.7	709	19 W30685	Glycerol kinase GU
5	26	89.7	709	19 W60262	Klebsiella pneumon
6	26	89.7	709	19 W57329	Mitochondrial glyc
7	26	89.7	709	20 Y26171	Cytosolic glycerol
8	26	89.7	766	20 Y06426	Avian topoisomeras
9	26	89.7	924	21 Y52991	Brosophila melanog
10	25	86.2	67	20 Y42644	Brassica napus D22
11	25	86.2	112	11 R05027	Deleted form of be
12	25	86.2	115	12 R14771	Truncated beta sub

13	25	86.2	187	20	Y02844	Fragment of human
14	25	86.2	218	18	W37339	Mouse HPRT. Mus m
15	25	86.2	243	11	R05026	Beta subunit of ra
16	25	86.2	244	19	W77619	Mercuric reductase
17	25	86.2	344	19	W63042	ATDMC1 protein seq
18	25	86.2	417	19	W41941	A. thaliana trunca
19	25	86.2	444	19	W47364	Arabidopsis thalia
20	25	86.2	445	18	W31506	Arabidopsis p-hydr
21	25	86.2	445	19	W41942	Correct A. thalian
22	25	86.2	445	20	W95708	Homo sapiens fetal
23	25	86.2	496	20	Y15820	Hydroxyphenyl pyru
24	25	86.2	531	16	R79382	Mouse CCT-zeta sub
25	25	86.2	652	18	W18010	Plasmodium falcipa
26	25	86.2	1125	19	W52288	Rattus norvegicus
27	25	86.2	1139	19	W37779	Rattus norvegicus
28	25	86.2	1242	19	W52287	Rattus norvegicus
29	24	82.8	9	18	W15161	Chlamydia trachoma
30	24	82.8	9	18	W15162	Chlamydia trachoma
31	24	82.8	9	18	W15169	Chlamydia trachoma
32	24	82.8	12	18	W25200	KOAGDV-peptide cap
33	24	82.8	12	20	W82483	X. riobravix toxin
34	24	82.8	14	16	R66361	Chlamydial MOMP VD
35	24	82.8	15	20	W74017	C. trachomatis MOM
36	24	82.8	17	11	R02115	Major Outer Membra
37	24	82.8	17	14	R37210	B-cell neutralisin
38	24	82.8	17	17	R95924	C. trachomatis maj
39	24	82.8	17	20	W84472	Peptide Ct4A deriv
40	24	82.8	17	20	W84545	Peptide 4A derived
41	24	82.8	19	15	R45873	Chlamydia trachoma
42	24	82.8	19	16	R66366	Hybrid capsid prot
43	24	82.8	19	16	R66376	Chlamydial MOMP VD
44	24	82.8	19	20	W84474	Peptide Ct4C deriv
45	24	82.8	19	20	W84547	Peptide 4C derived

ALIGNMENTS

RESULT 1	
W12695	
ID	W12695 standard; Protein; 541 AA.
XX	
AC	W12695;
XX	
DT	31-MAY-1997 (first entry)
XX	
DE	G-protein parathyroid hormone receptor HLTG74.
XX	
KW	G-protein parathyroid hormone receptor; HLTG74; parathormone; PTH;
KW	calcium; signal transduction; agonist; antagonist; hypocalcaemia;
KW	hyperphosphataemia; hypoparathyroidism; chronic tetany;
KW	osteoporosis; hypercalcaemia; hypoparathyroidism; hypophosphataemia;
KW	kidney stone; nephrolithiasis; therapy; diagnosis.
OS	Homo sapiens.
XX	
PN	W09639433-A1.
XX	
PD	12-DEC-1996.
XX	
PF	05-JUN-1995; 95WO-US07085.
XX	
PR	05-JUN-1995; 95WO-US07085.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Li Y, Rosen CA, Ruben SM, Soppet DR;
XX	
DR	WPI; 1997-043068/04.
DR	N-ESDB; T59619.
XX	
PT	Human G-protein parathyroid hormone receptor, HLTG74 - used to identify (ant)agonists, used in the treatment of hypo- or

SQ SEQUENCE 350 AA; 40315 MW; FC116B08392E4CFE CRC64;

Query Match 61.9%; Score 39; DB 5; Length 350;
Best Local Similarity 70.0%; Pred. No. 40;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 DSDSGTITIE 13
II : III : I :

Db 280 DSDGMITLED 289

RESULT 14

Q9RY57 PRELIMINARY; PRT; 742 AA.

AC Q9RY57;

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)

DE OXIDOREDUCTASE.

GN DRA0231.

OS Deinococcus radiodurans.

OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.

RP SEQUENCE FROM N.A.

RC STRAIN=RI;

RX MEDLINE; 20036896.

RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,

RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,

RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,

RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,

RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,

RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,

RA Fraser C.M.;

RT "Genome Sequence of the Radioresistant Bacterium Deinococcus

RT radiodurans RL";

RL Science 286:1571-1577(1999).

DR EMBL; AE001863; AAF12405.1; -.

DR TIGR; DRA0231; -.

DR INTERPRO; IPR000674; -.

DR INTERPRO; IPR000719; -.

DR PFAM; PF01315; Ald_xan_dh.C; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.

SQ SEQUENCE 742 AA; 79507 MW; 71BBD3F2BCC9B807 CRC64;

Query Match 61.9%; Score 39; DB 2; Length 742;
Best Local Similarity 54.3%; Pred. No. 93;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 QLSDSGTITIE 12

II : III : I :

Db 458 RLEQDGTITVE 468

RESULT 15

Q9VW06

ID Q9VW06 PRELIMINARY; PRT; 1339 AA.

AC Q9VW06;

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)

DE CG9279 PROTEIN.

GN CG9279.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY.

RX MEDLINE; 20196006.

RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

DR EMBL; AE003516; AAF49149.1; -.

DR FLYBASE; FBgn0036882; CG9279.

DR INTERPRO; IPR000938; -.

DR PFAM; PF01302; CAP_GLY; 1.

DR PROSITE; PS00845; CAP_GLY; 1.

SQ SEQUENCE 1339 AA; 152535 MW; 3BB624E7EBC7C0B6 CRC64;

Query Match 61.9%; Score 39; DB 5; Length 1339;
Best Local Similarity 58.3%; Pred. No. 1.8e+02;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 QLSDSGTITIE 13

II : III : I :

Db 693 QL RADGTLSTIED 704

Search completed: November 8, 2000, 08:56:01
Job time: 512 sec

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "the genome sequence of *Brosophila melanogaster*.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003753; AAF56519.1; -;
DR FLYBASE; FBgn0039380; CG5890.
DR INTERPRO; IPR001125; -;
DR INTERPRO; IPR002048; -;
DR PRAM; PF00036; efnhand; 3.
DR PRINTS; PR00450; RECOVERIN.
DR PROSITE; PS00018; EF_HAND; 3.
SQ SEQUENCE 206 AA; 23972 MW; 4699B324A54DD765 CRC64;

Query Match 61.9%; Score 39; DB 5; Length 206;
Best Local Similarity 72.7%; Pred. No. 22;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 LQDGGITIEE 13
I: || |||||
Db 174 LQDGGITIEE 184

RESULT 12

P95136 PRELIMINARY; PRT; 243 AA.

AC P95136; Q50460;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE HYPOTHETICAL 26.5 KDA PROTEIN.
GN RV2956 OR MTCY349.33C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE; 98295987.
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churche C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaita F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagsis K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of *Mycobacterium tuberculosis* from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
[2]
RP SEQUENCE FROM N.A.
RA Smith D.R.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Robison K.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z83018; CAB05420.1; -;
DR EMBL; U00024; AAA50939.1; -;
DR TUBERCULIST; RV2956; -;
KW Hypothetical protein.
SQ SEQUENCE 243 AA; 26566 MW; D9F196111185B501A CRC64;

Query Match 61.9%; Score 39; DB 2; Length 243;
Best Local Similarity 87.5%; Pred. No. 27;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 DSDGTITI 11
I: |||||I|
Db 98 DSDGTITI 105

RESULT 13

Q9VCQ3 PRELIMINARY; PRT; 350 AA.

AC Q9VCQ3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CG4624 PROTEIN.
GN CG4624.
OS *Drosophila melanogaster* (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE; 20196006.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhattacharya S., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "the genome sequence of *Brosophila melanogaster*.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003742; AAF56104.1; -;
DR FLYBASE; FBgn0039058; CG4624.

Query Match 61.9%; Score 39; DB 10; Length 151;
 Best Local Similarity 70.0%; Pred. No. 16;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 DSDGTITIEE 13
 I I I I I I I I
 Db 24 DGDGCITVEE 33

RESULT 9
 Q90UX7 PRELIMINARY; PRT; 175 AA.
 AC Q90UX7;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
 DE CALCINEURIN SUBUNIT.
 GN L7171.06.
 OS Leishmania major.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FRIEDLIN;
 RA Tosato V., Bruschi C.V., Ivens A.C., Murphy L., Quail M.,
 RA Rajandream M.A., Barrell B.G.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FRIEDLIN;
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
 RA Smith D.F.;
 RT "A physical map of the Leishmania major Friedlin genome.";
 RL Genome Res. 8:135-145(1998).
 DR EMBL: AL133435; CAB62809.1; -;
 DR INTERPRO: IPR001125; -;
 DR INTERPRO: IPR002048; -;
 DR PFAM: PF00036; ehand; 4.
 DR PRINTS: PR00450; RECOVERIN.
 DR PROSITE: PS00018; EF_HAND; UNKNOWN.2.
 SQ SEQUENCE 175 AA; 13660 MW; 9448F127F4DFA0EB CRC64;

Query Match 61.9%; Score 39; DB 5; Length 175;
 Best Local Similarity 61.5%; Pred. No. 19;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AQLDSDGTITIEE 13
 I I I I I I I I
 Db 145 ADVDRDGYITFEE 157

RESULT 10
 Q18282 PRELIMINARY; PRT; 193 AA.
 AC Q18282;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
 DE CODED FOR BY C. ELEGANS CDNA CEMS20R.
 GN C28C12.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE; 94150718.
 RA Wilton R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favellio A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,

RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smailon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RL Nature 368:32-38(1994).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Miller N.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U40797; AAB37545.1; -;
 SQ SEQUENCE 193 AA; 21752 MW; 2B8E7EEF5F46712B CRC64;

Query Match 61.9%; Score 39; DB 5; Length 193;
 Best Local Similarity 70.0%; Pred. No. 21;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 QLQSDGTITTI 11
 I I I I I I I I
 Db 107 QLQSDGELTL 116

RESULT 11
 Q9VBL2 PRELIMINARY; PRT; 206 AA.
 AC Q9VBL2;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
 DE CG5890 PROTEIN.
 GN CG5890.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE; 20196006.
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Adair J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Beriman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

```

GN F20B18.80.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Rosidae; eurosids II; Brassicales;
OC Brassicaceae; Arabidopsi[s].
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Rose M., Hempel S., Entian K.-D., Hoheisel J., Mewes H.W.,
RA Mayer K.F.X., Schueller C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL049483; CAB39662.1; -.
DR INTERPRO: IPR002048; -.
DR PFAM: PF000036; ehand; 2.
DR PROSITE: PS00018; EF_HAND; UNKNOWN_2.
SQ SEQUENCE 628 AA; 69751 MW; 24234F352207F5CC CRC64;

Query Match 63.5%; Score 40; DB 10; Length 628;
Best Local Similarity 53.8%; Pred. No. 52;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AQLSDSGTITIEE 13
I : : : : :
Db 203 ADLNCGVVTTIDE 215

RESULT 6
Q9X616
ID Q9X616 PRELIMINARY; PRT; 74 AA.
AC Q9X616;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE GROEL PROTEIN (FRAGMENT).
CN GROEL.
OS Chloroflexus aurantiacus.
OC Bacteria; Green non-sulfur bacteria; Chloroflexaceae group;
OC Chloroflexaceae; Chloroflexus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 99058143.
RA Gupta R.S.;
RT "Protein phylogenies and signature sequences: A reappraisal of
RT evolutionary relationships among archaeobacteria, eubacteria, and
RT eukaryotes.";
RL Microbiol. Mol. Biol. Rev. 62:1435-1491(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Gupta R.S., Mukhtar T., Singh B.;
RT "Evolutionary relationships among photosynthetic prokaryotes
RT (Helicobacterium chlorum, Chloroflexus auranticus, Cyanobacteria,
RT Chlorobium tepidum and Proteobacteria): implications regarding the
RT origin of photosynthesis.";
RL Mol. Microbiol. 0:0-0(1999).
DR EMBL: AF130450; AAD33468.1; -.
DR INTERPRO: IPR002423; -.
DR PFAM: PF00118; cpn60_TCP1; 1.
FT NON_TER 1
FT TER 74
FT SEQUENCE 74 AA; 7498 MW; 3514B94063E6D14E CRC64;

Query Match 61.9%; Score 39; DB 2; Length 74;
Best Local Similarity 53.8%; Pred. No. 7.1;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AQLSDSGTITIEE 13
I : : : : :
Db 60 ARVKGKGVITVEE 72

```

```

RESULT 7
Q39890
ID Q39890 PRELIMINARY; PRT; 150 AA.
AC Q39890;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CALMODULIN.
RN SCAM-4.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Fabales; Fabaceae;
OC Papilionoideae; Glycine.
RN [1]
RP SEQUENCE FROM N.A.
RA Lee S.H., Kim J.C., Choi Y.J., Bahk J.D., Hong J.C., Cho M.J.;
RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL: L01433; AAA34015.1; -.
DR HSP: P02593; ITRC.
DR MENDEL: 8263; Glyma;1222;8263.
DR INTERPRO: IPR002048; -.
DR PFAM: PF00036; ehand; 4.
DR PROSITE: PS00018; EF_HAND; UNKNOWN_4.
SQ SEQUENCE 150 AA; 17018 MW; E074D5E60F89893C CRC64;

Query Match 61.9%; Score 39; DB 10; Length 150;
Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 DSDGTITIEE 13
I : : : : :
Db 23 DGDGCITVEE 32

RESULT 8
Q23320
ID Q23320 PRELIMINARY; PRT; 151 AA.
AC Q23320;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CALMODULIN.
CN CAM8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Rosidae; eurosids II; Brassicales;
OC Brassicaceae; Arabidopsi[s].
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., Terryn N.,
RA Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,
RA Puigdomenech P., Hatzopoulos P., Obermaier B., Duesterhoft A.,
RA Jones J., Palme K., Ansoorge W., Delseny M., Bancroft I., Mewes H.W.,
RA Schueller C., Chalvatzis N.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Zielinski R.E.;
RT "Identification of three new sequences encoding conserved and highly
RT divergent calmodulin isoforms from Arabidopsis thaliana.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z97336; CAB10243.1; -.
DR EMBL: AF178074; AAD53314.1; -.
DR HSP: P02593; ITRC.
DR MENDEL: 26687; Arabid;1222;26687.
DR INTERPRO: IPR002048; -.
DR PFAM: PF00036; ehand; 4.
DR PROSITE: PS00018; EF_HAND; UNKNOWN_4.
SQ SEQUENCE 151 AA; 17156 MW; A72571DC7EAB2427 CRC64;

```

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 QLSDSGTTTIEE 13
|:|||||:|
Db 48 QLSDSPTLTIAE 59

RESULT 2

Q18302 ID O18302 PRELIMINARY; PRT; 390 AA.
AC O18302;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JUN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
DE ZK849.2 PROTEIN.
CN ZK849.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA Kershaw J.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94150718.
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
Smaalen N., Smith A., Sonnhammer E., Staden R., Sulston J.,
Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
Watson A., Weinstock L., Wilkinson-Sprout J., Wohlman P.,
"2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans.";
RL Nature 368:32-38(1994).
DR EMBL: 282095; CAB05026.1; -.
DR HSSP: Q12959; IPDR.
DR INTERPRO: IPR000408; -.
DR INTERPRO: IPR001478; -.
DR PFAM: PF00595; PD2; 1.
DR PROSITE: PS00626; RCL1.2; UNKNOWN_1.
SQ SEQUENCE 390 AA; 43612 MW; F3DDE3394BBF631C CRC64;

Query Match 63.5%; Score 40; DB 5; Length 390;
Best Local Similarity 77.8%; Pred. No. 30;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 DSDSGTTTIE 12
|:|||||
Db 376 DDDGTVTIE 384

RESULT 3

Q57002 ID Q57002 PRELIMINARY; PRT; 543 AA.
AC Q57002;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE CHAPERONIN LIKE PROTEIN.
GN GROEL2.
OS Synchococcus vulcanus.
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 96248387.
RA Furuki M., Tanaka N., Hiyama T., Nakamoto H.;
RT "Cloning, characterization and functional analysis of groEL-like gene

RT from thermophilic cyanobacterium Synchococcus vulcanus, which does
RT not form an operon with groES.";
RL Biochim. Biophys. Acta 1294:106-110(1996).
DR EMBL: D86384; BAAL3082.1; -.
DR HSSP: P06139; IGRLL.
DR INTERPRO: IPR001844; -.
DR INTERPRO: IPR002423; -.
DR PFAM: PF00118; cpn60_TCP1; 1.
DR PRINTS: PR00298; CHAPERONIN60.
DR PRINTS: PR00304; TCOMPLEXTCP1.
SQ SEQUENCE 543 AA; 57102 MW; 65171A43BDICE3E1 CRC64;

Query Match 63.5%; Score 40; DB 2; Length 543;
Best Local Similarity 61.5%; Pred. No. 44;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 AOLDSGTTTIEE 13
|:|:|:|:|
Db 165 AKVGKDGVTIEE 177

RESULT 4

Q9SDM4 ID Q9SDM4 PRELIMINARY; PRT; 595 AA.
AC Q9SDM4;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
DE CALCIUM-DEPENDENT PROTEIN KINASE.
OS Dunalialia tertiolecta.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Dunalialaceae; Dunalialia.
RN [1]
RP SEQUENCE FROM N.A.
RA Pinontoan R., Yuasa T., Anderca M.I., Matsuoaka T., Uozumi N., Mori H.,
Muto S.;
RT "Cloning a CDPK from Dunalialia tertiolecta.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF216527; AAF21062.1; -.
DR INTERPRO: IPR000008; -.
DR INTERPRO: IPR000719; -.
DR INTERPRO: IPR002048; -.
DR INTERPRO: IPR002290; -.
DR PFAM: PF00036; ehand; 4.
DR PFAM: PF00069; pkinase; 1.
DR PFAM: PF00168; C2; 1.
DR PROSITE: PS00018; EF_HAND; UNKNOWN_4.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50004; C2_DOMAIN_2; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Kinase.
SQ SEQUENCE 595 AA; 65749 MW; 5C20DE30027A6CFE CRC64;

Query Match 63.5%; Score 40; DB 10; Length 595;
Best Local Similarity 61.5%; Pred. No. 49;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AOLDSGTTTIEE 13
|:|:|:|:|
Db 504 ADVGKDGTTIDYER 516

RESULT 5

Q9SZH1 ID Q9SZH1 PRELIMINARY; PRT; 628 AA.
AC Q9SZH1;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
DE PUTATIVE PHOSPHATIDYLSEINE DECARBOXYLASE.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: November 8, 2000, 08:55:58 ; Search time 152.43 Seconds
(without alignments)
7.963 Million cell updates/sec

Title: US-09-236-468a-2_COPY_24_36
Perfect score: 63
Sequence: 1 AQLSDSGTTIEE I3

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 297973

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: SP_TREMBL_14.*
- 2: sp_archaea.*
- 3: sp_bacteria.*
- 4: sp_fungi.*
- 5: sp_human.*
- 6: sp_invertebrate.*
- 7: sp_mammal.*
- 8: sp_mhc.*
- 9: sp_organelle.*
- 10: sp_phage.*
- 11: sp_plant.*
- 12: sp_rodent.*
- 13: sp_virus.*
- 14: sp_vertebrate.*
- 15: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	65.1	195	5 Q9XTN7	Q9xtn7 caenorhabdi
2	40	63.5	390	5 O18302	O18302 caenorhabdi
3	40	63.5	543	2 Q57002	Q57002 synchococc
4	40	63.5	595	10 Q9SDM4	Q9sdm4 dunaliella
5	40	63.5	628	10 Q9SZH1	Q9szhl arabidopsis
6	39	61.9	74	2 Q9X616	Q9x616 chloroflexu
7	39	61.9	150	10 Q39890	Q39890 glycine max
8	39	61.9	151	10 Q23320	Q23320 arabidopsis
9	39	61.9	175	5 Q9U0X7	Q9u0x7 leishmania
10	39	61.9	193	5 Q18282	Q18282 caenorhabdi
11	39	61.9	206	5 Q9VBL2	Q9vbl2 drosophila
12	39	61.9	243	2 P95136	P95136 mycobacteri
13	39	61.9	350	5 Q9VCQ3	Q9vcq3 drosophila
14	39	61.9	742	2 Q9RYS7	Q9rys7 deinococcus
15	39	61.9	1339	5 Q9VM06	Q9vm06 drosophila
16	38	60.3	147	5 Q9XZP3	Q9xzp3 brachiosteo
17	38	60.3	150	10 Q40982	Q40982 pisum sativ
18	38	60.3	169	10 Q2845	Q2845 arabidopsis
19	38	60.3	209	10 Q9SS31	Q9ss31 arabidopsis

20	38	60.3	218	2	O53558	O53558 mycobacteri
21	38	60.3	453	10	Q9S9V0	Q9s9v0 arabidopsis
22	38	60.3	494	10	Q9ZSA4	Q9zsa4 arabidopsis
23	38	60.3	548	2	O53560	O53560 mycobacteri
24	38	60.3	571	10	O48827	O48827 arabidopsis
25	38	60.3	702	2	O83856	O83856 treponema p
26	38	60.3	742	4	Q9UQ29	Q9uq29 homo sapien
27	38	60.3	884	2	Q9RW10	Q9rw10 deinococcus
28	38	60.3	910	5	Q9U3F8	Q9u3f8 caenorhabdi
29	38	60.3	924	5	Q20239	Q20239 caenorhabdi
30	38	60.3	2047	4	Q9UH61	Q9uh61 homo sapien
31	38	60.3	2168	11	O61043	O61043 mus musculu
32	37	58.7	71	5	Q9V687	Q9v687 drosophila
33	37	58.7	120	6	Q29376	Q29376 sus scrofa
34	37	58.7	136	13	Q91972	Q91972 oryzias lat
35	37	58.7	138	5	O15931	O15931 symbiodiniu
36	37	58.7	140	5	O25420	O25420 leishmania
37	37	58.7	146	5	Q94801	Q94801 toxoplasma
38	37	58.7	149	3	Q94739	Q94739 pleurotosi o
39	37	58.7	149	5	O00897	O00897 dictyosteli
40	37	58.7	149	5	O02367	O02367 ciona inte
41	37	58.7	149	5	O96081	O96081 halocynthia
42	37	58.7	149	5	O97341	O97341 suberites d
43	37	58.7	149	5	O16305	O16305 caenorhabdi
44	37	58.7	149	5	Q9V3T4	Q9v3t4 drosophila
45	37	58.7	149	13	O93410	O93410 gallus gall

ALIGNMENTS

RESULT 1
Q9XTN7 PRELIMINARY; PRT; 195 AA.
AC Q9XTN7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE C32H11.11 PROTEIN.
GN C32H11.11
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
RA Barlow K.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.";
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94150718.

RA Wilson R., Alnscough R., Anderson K., Baynes C., Berks M., Berts M., Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A., Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirsten J., Laister N., Latreille P., Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R., Stalson N., Smith A., Sonnhammer E., Staden K., Waterston R., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.";
RL Nature 368:32-38(1994).
DR EMBL; Z82260; CAB05137.1; -;
DR EMBL; Z82260; CAB05134.1; -;
SQ SEQUENCE 195 AA; 21914 MW; 10B9FB636D932B3C CRC64;

Query Match 65.1%; Score 41; DB 5; Length 195;
Best Local Similarity 66.7%; Pred. No. 9.3;

OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 88163497.
 RA Mangelndorf D.J., Komm B.S., McDonnell D.P., Pike J.W., Haussler M.R.;
 RT "Immunoselection of cDNAs to avian intestinal calcium binding protein
 RT 28K and a novel calmodulin-like protein: assessment of mRNA
 RT regulation by the vitamin D hormone.";
 RL Biochemistry 26:8332-8338(1987).
 CC -!- SIMILARITY: THIS PROTEIN SEEMS TO DIFFER FROM CALMODULIN BY A
 CC SINGLE POSITION (S-86 INSTEAD OF G-86).
 CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M18355; AAA48645.1; -.
 DR PIR; A29422; A29422.
 DR HSP; P02593; IAK8.
 DR INTERPRO; IPR002048; -.
 DR PFAM; PF00036; efhand; 4.
 DR PROSITE; PS00018; EF_HAND; 4.
 KW Calcium-binding; Duplication.
 FT NON_TER 1
 FT CA_BIND 10 21 SITE 1.
 FT CA_BIND 46 57 SITE 2.
 FT CA_BIND 83 94 SITE 3.
 FT CA_BIND 119 130 SITE 4.
 FT NON_TER 131 131
 SQ SEQUENCE 131 AA; 14847 MW; 3BF82910B8CA3425 CRC64;

Query Match 58.7% Score 37; DB 1; Length 131;
 Best Local Similarity 70.0%; Pred. No. 15;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 DSGDTITTEE 13
 Db 12 DGDGTTTKE 21

Search completed: November 8, 2000, 09:03:42
 Job time: 852 sec

"Touch-inducible genes for calmodulin and a calmodulin-related protein are located in tandem on a chromosome of Arabidopsis thaliana.";
 RL Plant Cell Physiol. 36:1369-1373(1995).
 [3]

RP SEQUENCE OF 3-74 FROM N.A.

RC STRAIN=CV. COLUMBIA;

RX MEDLINE; 90150263.

RA Braam J., Davis R.W.;

RT "Rain", wind-, and touch-induced expression of calmodulin and

calmodulin-related genes in Arabidopsis.";

RL Cell 60:357-364(1990).

CC -1- FUNCTION: BINDS CALCIUM.

CC -1- INDUCTION: BY RAIN-, WIND-, AND TOUCH (THIGMOMORPHOGENESIS).

CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; L34546; AAC37419.1; -.

DR EMBL; D45848; BAA08282.1; -.

DR PIR; B34669; B34669.

DR HSSP; P02593; LCGL.

DR INTERPRO; IPR002048; -.

DR PFAM; PF00036; ehand; 6.

DR PROSITE; PS00018; EF_HAND; 6.

KW Calcium-binding; Repeat.

FT CA_BIND 21 32

FT CA_BIND 57 68 1 (POTENTIAL).

FT CA_BIND 110 121 2 (POTENTIAL).

FT CA_BIND 146 157 3 (POTENTIAL).

FT CA_BIND 200 211 4 (POTENTIAL).

FT CA_BIND 236 247 5 (POTENTIAL).

FT CA_BIND 267 278 6 (POTENTIAL).

FT CONFLICT 14 14 R -> K (IN REF. 3).

FT CONFLICT 36 36 M -> V (IN REF. 3).

FT CONFLICT 40 43 IGEK -> LGON (IN REF. 3).

FT CONFLICT 52 52 L -> M (IN REF. 3).

FT CONFLICT 56 56 A -> V (IN REF. 3).

FT CONFLICT 71 72 CV -> NL (IN REF. 3).

FT CONFLICT 223 226 KQKL -> OTKA (IN REF. 2).

SQ SEQUENCE 324 AA; 36876 MW; 5978AECED9BCC7 CRC64;

Query Match 60.3%; Score 38; DB 1; Length 324;

Best Local Similarity 61.5%; Pred. No. 26;

Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AQLSDSGTITIEE 13

I I I I I I I

Db 56 ADLDGDGTIDPE 68

RESULT 14

ID KQGB_RAT

AC P49621;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE DIACYLGLYCEROL KINASE, BETA (EC 2.7.1.107) (DIGLYCERIDE KINASE) (DGK-

BETA) (DAG KINASE BETA) (90 KDA DIACYLGLYCEROL KINASE).

GN DAGK2.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NC [1]

RP SEQUENCE FROM N.A.

RC STRAIN=WISTAR; TISSUE=BRAIN;

RX MEDLINE; 93361494.
 RA Goto K., Kondo H.;
 RT "Molecular cloning and expression of a 90-kDa diacylglycerol kinase
 that predominantly localizes in neurons.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:7598-7602(1993).
 CC -1- FUNCTION: EXHIBIT HIGH PHOSPHORYLATION ACTIVITY FOR LONG-CHAIN
 DIACYLGLYCEROLS.
 CC -1- CATALYTIC ACTIVITY: ATP + 1,2-DIACYLGLYCEROL -> ADP +
 1,2-DIACYLGLYCEROL 3-PHOSPHATE.
 CC -1- ENZYME REGULATION: STIMULATED BY PHOSPHATIDYLSELINE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. CAN BE LOOSELY BOUND TO THE
 MEMBRANES.
 CC -1- TISSUE SPECIFICITY: CONFINED TO RESTRICTED NEURONAL POPULATIONS
 SUCH AS THE CAUDATE-PUTAMEN, THE ACCUMBENS NUCLEUS, AND THE
 OLFACTORY TUBERCLE.
 CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC DIACYLGLYCEROL KINASE
 FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
 BINDING DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL; D16100; BAA03675.1; -.

DR HSSP; P05697; ITBO.

DR INTERPRO; IPR000756; -.

DR INTERPRO; IPR001206; -.

DR INTERPRO; IPR002048; -.

DR INTERPRO; IPR002219; -.

DR PFAM; PF00609; DAGKa; 1.

DR PFAM; PF00781; DAGKc; 1.

DR PFAM; PF00130; DAG_PE-bind; 2.

DR PFAM; PF00036; ehand; 2.

DR PRINTS; PR00008; DAGPEDOMAIN.

DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.

DR PROSITE; PS00081; DAG_PE_BIND_DOM_2; 2.

DR PROSITE; PS00018; EF_HAND; 2.

KW Transferase; Kinase; Calcium-binding; Phorbol-ester binding;

Multigene family.

FT CA_BIND 161 172 SITE 1 (POTENTIAL).

FT CA_BIND 206 217 SITE 2 (POTENTIAL).

FT DOMAIN 244 292 PHORBOL-ESTER AND DAG BINDING (BY

FT DOMAIN 308 356 SIMILARITY).

FT DOMAIN 433 559 PHORBOL-ESTER AND DAG BINDING (BY

FT DOMAIN 579 759 SIMILARITY).

FT CATALYTIC-A (POTENTIAL).

FT CATALYTIC-B (POTENTIAL).

SQ SEQUENCE 801 AA; 90288 MW; F30874CD2DCE363D CRC64;

Query Match 60.3%; Score 38; DB 1; Length 801;

Best Local Similarity 60.0%; Pred. No. 67;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 DSDGTITIEE 13

I I I I I I I

Db 208 DRDGTVSLEE 217

RESULT 15

CALN_CHICK

ID CALN_CHICK

AC P05419;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-NOV-1988 (Rel. 09, Last sequence update)

DT 01-JAN-1990 (Rel. 13, Last annotation update)

DE NEO-CALMODULIN (NEOCAM) (FRAGMENT).

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D89970; BAA14046.1; -
 DR INTERPRO: IPR001844; -
 DR INTERPRO: IPR002423; -
 DR PFAM: PF00118; cpn60_TCPL1; 1.
 DR PRINTS: PR00298; CHAPERONIN60.
 DR PRINTS: PR00304; TCOMPLEXCP1.
 DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
 KW Chaperone; ATP-binding.
 SQ SEQUENCE 554 AA; 59158 MW; 5CCB9866F5AA2F08 CRC64;

Query Match 61.9%; Score 39; DB 1; Length 554;
 Best Local Similarity 58.3%; Pred. No. 31;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QLSDSGTITIEE 13
 : : ||| ||| :
 Db 169 KVGSDGVITVEE 180

RESULT 11
 ID E631_DROME STANDARD; PRT; 193 AA.
 AC P48593;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CALCIUM-BINDING PROTEIN E63-1.
 GN EIP63F-1 OR E63-1.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CANTON-S;
 RX MEDLINE; 95401881.
 RA Andres A.J., Thummel C.S.;
 RT "The Drosophila 63F early puff contains E63-1, an ecdysone-inducible
 RT gene that encodes a novel Ca(2+)-binding protein.";
 RL Development 121:2667-2679(1995).
 CC -!- INDUCTION: BY ECDYSONE.
 CC -!- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U25882; AAB61120.1; -
 DR HSP: P02593; ITRC.
 DR FLYBASE: FBgn0004910; Eip63F-1.
 DR INTERPRO: IPR002048; -
 DR PFAM: PF00036; ehand; 3.
 DR PROSITE: PS00018; EF_HAND; 3.
 KW Calcium-binding; Repeat.
 FT CA_BIND 48 59 SITE 1 (POTENTIAL).
 FT CA_BIND 98 109 ANCESTRAL CALCIUM SITE 2 (POTENTIAL).
 FT CA_BIND 140 151 SITE 3 (POTENTIAL).
 FT CA_BIND 176 187 SITE 4 (POTENTIAL).
 SQ SEQUENCE 193 AA; 22013 MW; D9C117B9447274A3 CRC64;

Query Match 60.3%; Score 38; DB 1; Length 193;
 Best Local Similarity 61.5%; Pred. No. 15;

Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 AQLSDSGTITIEE 13
 : : ||| ||| :
 Db 175 ADLQDQGRINVEE 187

RESULT 12
 ID YNB0_YEAST STANDARD; PRT; 241 AA.
 AC P53981;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE HYPOTHETICAL 27.5 KDA PROTEIN IN SPOL-SIS1 INTERGENIC REGION.
 GN YNL010W OR N2866.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomycetes.

RN [1]
 RP SEQUENCE FROM N.A.
 RA Andre B., Iraqui Houssaini I., Urrestarazu L.A., Visiers S.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Z71286; CAA95870.1; -
 DR SGD: S0004955; YNL010W.
 KW Hypothetical protein.
 SQ SEQUENCE 241 AA; 27480 MW; 38F5D7FF14018637 CRC64;

Query Match 60.3%; Score 38; DB 1; Length 241;
 Best Local Similarity 60.0%; Pred. No. 19;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 DSDGTITIEE 13
 : : ||| : :
 Db 9 DFDGTVTLED 18

RESULT 13
 ID TCH3_ARATH STANDARD; PRT; 324 AA.
 AC P25071; Q38972; Q39064;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE CALMODULIN-RELATED PROTEIN 3, TOUCH-INDUCED.
 GN TCH3 OR CAL4.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RX MEDLINE; 95128169.
 RA Sistrunk M.L., Antosiewicz D.M., Purugganan M.M., Braam J.;
 RT "Arabidopsis TCH3 encodes a novel Ca2+ binding protein and shows
 RT environmentally induced and tissue-specific regulation.";
 RL Plant Cell 6:1553-1565(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, LANDSBERG ERECTA;
 RX MEDLINE; 96104307.
 RA Ito T., Hirano M., Akama K., Shimura Y., Okada K.;

```
OY 1 AQLSDSGTITIEE 13
  ||| ||| |||
Db 143 ADDKDKGISFEE 155

RESULT 8
Y443_METJA
ID Y443_METJA STANDARD; PRT; 227 AA.
AC Q57885;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL PROTEIN MJ0443.
GN MJ0443.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
CC Methanococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE; 96337999.
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: SOME. TO S.CEREVISIAE YCL59C.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U67495; AAB98430.1; -.
DR HSSP; P02633; 2BCB.
DR TIGR; MJ0443; -.
DR INTERPRO; IPR000958; -.
DR PFAM; PF00013; KH-domain; 2.
DR Hypothetical protein.
SQ SEQUENCE 227 AA; 25839 MW; 9E59A5FE3CDBE89C CRC64;

Query Match 63.5%; Score 40; DB 1; Length 227;
Best Local Similarity 60.0%; Pred. No. 8.4;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 OLDSDGTITI 11
  ::|||::||
Db 49 EIDADCTVTI 58

RESULT 9
CH60_LACLA
ID CH60_LACLA STANDARD; PRT; 542 AA.
AC P37282;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN).
GN MOPA OR GROEL.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
RN [1]

SEQUENCE FROM N.A.
MEDLINE; 93252268.
Kim S.G., Batt C.A.;
"Cloning and sequencing of the Lactococcus lactis subsp. lactis
groEL operon.";
Gene 127:121-126(1993).
-1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CONDITIONS (BY SIMILARITY).
-1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
7 SUBUNITS (BY SIMILARITY).
-1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X71132; CAA50446.1; -.
DR PIR; S32106; S32106.
DR PIR; JN0661; JN0661.
DR HSSP; P06139; 1GRL.
DR INTERPRO; IPR001844; -.
DR INTERPRO; IPR002423; -.
DR PFAM; PF00118; cpn60_TCP1; 1.
DR PRINTS; PR00298; CHAPERONIN60.
DR PRINTS; PR00304; TCOMPLEXTCP1.
DR PROSITE; PS00296; CHAPERONINS_CPN60; 1.
KW Chaperone; ATP-binding.
SQ SEQUENCE 542 AA; 57188 MW; 26F2326F3F34663A CRC64;

Query Match 63.5%; Score 40; DB 1; Length 542;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 OLDSDGTITIEE 13
  ::||| ||| |||
Db 166 RVGSDGVITIEE 177

RESULT 10
CH60_HOLOB
ID CH60_HOLOB STANDARD; PRT; 554 AA.
AC P94820;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60 KDA CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN).
GN MOPA OR GROEL.
OS Holospora obtusa.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC rickettsia group ciliate endosymbionts; Holospora.
RN [1]
RP SEQUENCE FROM N.A.
RA Dohra H., Fujishima M., Ishikawa H.;
RT "Structure and expression of a groE operon homolog of the ciliate
RT macronucleus-s specific symbiont Holospora obtusa of the ciliate
RT Paramecium caudatum.";
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CONDITIONS (BY SIMILARITY).
-1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
7 SUBUNITS (BY SIMILARITY).
-1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```


DR EMBL; AL7083; CAA01320.1; -.
 DR PIR; JN0097; JN0097.
 KW SIGNAL.
 FT CHAIN 1 27
 FT CHAIN 28 461 SECRETED 45 KDA PROTEIN.
 FT DOMAIN 299 314 POLY-SER.
 SQ SEQUENCE 461 AA; 47387 MW; 51493C42224F3C03 CRC64;

Query Match 66.7%; Score 42; DB 1; Length 461;
 Best Local Similarity 72.7%; Pred. No. 8.2;
 Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LQSDGTITIEE 13
 Db 426 VNSDGTITIEE 436
 :|||||||:
 :|||||||:

RESULT 4
 GCP_MYCTU
 ID GCP_MYCTU STANDARD; PRT; 344 AA.
 AC Q50709;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROBABLE O-SIALOGLYCOPROTEIN ENDOPEPTIDASE (EC 3.4.24.57)
 DE (GLYCOPROTEASE).
 GN GCP OR RV3419C OR MTCY78.10.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE; 98295987.
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaita F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrett B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence.";
 RL Nature 393:537-544(1998).
 CC -1- FUNCTION: COULD BE A METALLOPROTEASE.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF O-SIALOGLYCOPROTEINS; CLEAVES
 CC 31-ARG-1-ASP-32 BOND IN GLYCOPHORIN A. DOES NOT CLEAVE
 CC UNGLYCOSYLATED PROTEINS, DESIALYLATED GLYCOPROTEINS OR
 CC GLYCOPROTEINS THAT ARE ONLY N-GLYCOSYLATED.
 CC -1- COFACTOR: ZINC (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M22; ALSO KNOWN AS THE
 CC GLYCOPROTEASE FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; Z77165; CAB01004.1; -.
 DR TUBERCULIST; RV3419C; -.
 DR INTERPRO; IPR000905; -.
 DR PFAM; PF00814; Peptidase_M22; 1.
 DR PRINTS; PR00789; OSIALOPTASE.
 DR PROSITE; PS01016; GLYCOPROTEASE; 1.
 KW Hydrolase: Metalloprotease; Zinc.
 FT METAL 114 114 ZINC (POTENTIAL).
 FT METAL 118 118 ZINC (POTENTIAL).
 SQ SEQUENCE 344 AA; 35091 MW; 032B3EB3862CB2A4 CRC64;

Query Match 65.1%; Score 41; DB 1; Length 344;
 Best Local Similarity 63.6%; Pred. No. 8.8;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AQLSDSGTITI 11
 Db 19 ARLDPDGTITL 29
 :||| |||:
 :||| |||:

RESULT 5
 MOP_DESGI
 ID MOP_DESGI STANDARD; PRT; 907 AA.
 AC Q46509;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE ALDEHYDE OXIDOREDUCTASE (EC 1.2.-.-) (MOLYBDENUM IRON SULFUR PROTEIN).
 GN MOP.
 OS Desulfovibrio gigas.
 OC Bacteria; Proteobacteria; delta subdivision; Desulfovibrio.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 27774;
 RX MEDLINE; 94192692.
 RA Thoenes U., Flores O.L., Neves A., Devreese B., van Beeumen J.J.,
 RA Huber R., Romao M.J., Legall J., Moura J.J.G., Rodrigues-Pousada C.;
 RT "Molecular cloning and sequence analysis of the gene of the
 RT molybdenum-containing aldehyde oxidoreductase of Desulfovibrio
 RT gigas. The deduced amino acid sequence shows similarity to xanthine
 RT dehydrogenase.";
 RL Eur. J. Biochem. 220:901-910(1994).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS).
 RC STRAIN=ATCC 27774;
 RX MEDLINE; 96072968.
 RA Romao M.J., Archer M., Moura I., Moura J.J., Legall J., Engh R.,
 RA Schneider M., Hof P., Huber R.;
 RT "Crystal structure of the xanthine oxidase-related aldehyde oxido-
 RT reductase from D. gigas.";
 RL Science 270:1170-1176(1995).
 CC -1- COFACTOR: FAD, MOLYBDOPTERIN, AND TWO 2FE-2S CLUSTERS.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SIMILARITY: TO 2FE-2S FERREDOXINS IN THE N-TERMINAL DOMAIN.
 CC -1- SIMILARITY: STRONG, TO EUKARYOTIC XANTHINE DEHYDROGENASE.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; X77222; CAA54439.1; -.
 DR PDB; 1ALO; 11-JAN-97.
 DR INTERPRO; IPR000564; -.
 DR INTERPRO; IPR000674; -.
 DR INTERPRO; IPR001041; -.
 DR INTERPRO; IPR002888; -.
 DR PFAM; PF01315; Ald_xan_dh_C; 1.
 DR PFAM; PF00111; fer2; 1.
 DR PFAM; PF01799; fer2; 1.
 DR PROSITE; PS00197; 2FE2S_FERREDOXIN; 1.
 KW Oxidoreductase; NAD; Molybdenum; Flavoprotein; FAD; Iron-sulfur;
 3D-structure.
 FT METAL 40 40 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 FT METAL 45 45 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 FT METAL 48 48 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 FT METAL 60 60 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 SQ SEQUENCE 907 AA; 97034 MW; 898E7EEF708A64DF CRC64;

```
FT TRANSMEM 274 295 4 (POTENTIAL).
FT DOMAIN 296 313 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 314 334 5 (POTENTIAL).
FT DOMAIN 335 361 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 362 380 6 (POTENTIAL).
FT DOMAIN 381 391 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 392 414 7 (POTENTIAL).
FT DOMAIN 415 546 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 546 AA; 61800 MW; 2825AE040313527 CRC64;

Query Match 100.0%; Score 63; DB 1; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQLSDSGTITIEE 13
Db 24 AQLSDSGTITIEE 36

RESULT 2
PTR2_HUMAN STANDARD; PRT; 550 AA.
AC P49190;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PARATHYROID HORMONE RECEPTOR PRECURSOR (PTH2 RECEPTOR).
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RT TISSUE=BRAIN;
RX MEDLINE; 95318121.
RA Usdin T.B., Gruber C., Bonner T.I.;
RT "Identification and functional expression of a receptor selectively
recognizing parathyroid hormone, the PTH2 receptor.";
RL J. Biol. Chem. 270:15455-15458(1995).
[2]
SEQUENCE OF 26-40 AND 306-550 FROM N.A.
RX MEDLINE; 97079671.
RA Usdin T.B., Modi W., Bonner T.I.;
RT "Assignment of the human PTH2 receptor gene (PTH2) to chromosome 2q33
by fluorescence in situ hybridization.";
RL Genomics 37:140-141(1996).
CC -1- FUNCTION: THIS IS A SPECIFIC RECEPTOR FOR PARATHYROID HORMONE. THE
ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
ADENYLYL CYCLASE.
CC -1- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN BRAIN AND PANCREAS.
CC ALSO EXPRESSED IN THE TESTIS.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; U25128; AAC50157.1; -
CC EMBL; U47124; AAA96796.1; -
CC EMBL; U47129; AAC50767.1; -
CC EMBL; U47125; AAC50767.1; JOINED.
CC EMBL; U47126; AAC50767.1; JOINED.
CC EMBL; U47127; AAC50767.1; JOINED.
CC EMBL; U47128; AAC50767.1; JOINED.
```

```
DR GCRDB; GCR_2003; -
DR MIM; 601469; -
DR INTERPRO; IPR000832; -
DR PFAM; PF00002; 7tm_2; 1.
DR PRINTS; PR00249; GPCRSECRETIN
DR PROSITE; PS00649; G-PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G-PROTEIN_RECEP_F2_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 550 PARATHYROID HORMONE RECEPTOR.
FT DOMAIN 27 145 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 146 169 1 (POTENTIAL).
FT DOMAIN 170 176 2 (POTENTIAL).
FT TRANSMEM 177 196 2 (POTENTIAL).
FT DOMAIN 197 237 3 (POTENTIAL).
FT TRANSMEM 238 260 3 (POTENTIAL).
FT DOMAIN 261 275 4 (POTENTIAL).
FT TRANSMEM 276 297 4 (POTENTIAL).
FT DOMAIN 298 317 5 (POTENTIAL).
FT TRANSMEM 317 337 5 (POTENTIAL).
FT DOMAIN 338 364 6 (POTENTIAL).
FT TRANSMEM 365 383 6 (POTENTIAL).
FT DOMAIN 384 394 7 (POTENTIAL).
FT TRANSMEM 395 417 7 (POTENTIAL).
FT DOMAIN 418 550 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 550 AA; 62235 MW; 2ADD14DBA68A9BF8 CRC64;

Query Match 100.0%; Score 63; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQLSDSGTITIEE 13
Db 24 AQLSDSGTITIEE 36

RESULT 3
US45_LACLC STANDARD; PRT; 461 AA.
AC P22865;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE SECRETED 45 KDA PROTEIN PRECURSOR.
OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
RN [1]
SEQUENCE FROM N.A., AND SEQUENCE OF 28-37.
RC STRAIN=MGI363;
RX MEDLINE; 91071599.
RA van Asseldonk M., Rutten G., Otteman M., Siezen R.J., de Vos W.M.,
RT "Cloning of usp45, a gene encoding a secreted protein from
Lactococcus lactis subsp. lactis MGI363.";
RL Gene 95:155-160(1990).
CC -1- SIMILARITY: TO E-FAECIUM P54.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; M60178; AAA25230.1; -
DR
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 09:03:41 : Search time 58.45 Seconds
(without alignments)
7.106 Million cell updates/sec

Title: US-09-236-468A-2_COPY_24_36
Perfect score: 63
Sequence: 1 AQLDSGTITIEE 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 87993

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	100.0	546	1 PTH2_RAT	P70555 rattus norv
2	63	100.0	550	1 PTR2_HUMAN	P49190 homo sapien
3	42	66.7	461	1 US45_LACLC	P22865 lactococcus
4	41	65.1	344	1 GCP_MYCTU	Q50709 mycobacteri
5	41	65.1	907	1 MOP_DESGI	O46509 desulfovibr
6	41	65.1	1043	1 SYI_METH	O27428 methanobact
7	40	63.5	174	1 CALB_NEUCR	P87072 neurospora
8	40	63.5	227	1 Y443_METJA	Q57885 methanococc
9	40	63.5	542	1 CH60_LACIA	P37282 lactococcus
10	39	61.9	554	1 CH60_HOLOB	P94820 holospira o
11	38	60.3	193	1 E631_DROME	P48593 drosophila
12	38	60.3	241	1 YNB0_YEAST	P53981 saccharomyc
13	38	60.3	324	1 TCH3_ARATH	P25071 arabidopsi
14	38	60.3	801	1 KDGB_RAT	P49621 rattus norv
15	37	58.7	131	1 CALN_CHICK	P05419 gallus gall
16	37	58.7	138	1 CALL_ARBP	P05932 arabacia pun
17	37	58.7	148	1 CALM_ACHKL	P15094 achlya kleb
18	37	58.7	148	1 CALM_DROME	P07181 drosophila
19	37	58.7	148	1 CALM_DELEL	P02594 electrophor
20	37	58.7	148	1 CALM_EUGGR	P11118 euglena gra
21	37	58.7	148	1 CALM_HUMAN	P02593 homo sapien
22	37	58.7	148	1 CALM_METSE	P02596 metridium s
23	37	58.7	148	1 CALM_PARTE	P07463 paramecium
24	37	58.7	148	1 CALM_PATSP	P02595 patinopecte
25	37	58.7	148	1 CALM_PHYIN	P27165 phytophthor
26	37	58.7	148	1 CALM_PLECO	P11120 pleurotus c
27	37	58.7	148	1 CALM_PLECO	P11121 pyuridae sp
28	37	58.7	148	1 CALM_STIJA	P21251 stichopus j
29	37	58.7	148	1 CALM_STYLE	P27166 stylonychia
30	37	58.7	148	1 CALM_TERYP	P02598 tetrahymena
31	37	58.7	148	1 CALM_TRYBB	P04465 trypanosoma
32	37	58.7	148	1 CALM_TRYCR	P18061 trypanosoma
33	37	58.7	149	1 CALM_PLAFA	P24044 plasmodium

RESULT 1

ID	PTH2_RAT	STANDARD	PRT	546 AA
AC	P70555			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	PARATHYROID HORMONE RECEPTOR PRECURSOR (PTH2 RECEPTOR).			
GN	PTH2.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 96426194.			
RA	Usdin T.B., Bonner T.I., Harta G., Mezey E.;			
RT	"Distribution of parathyroid hormone-2 receptor messenger ribonucleic acid in rat.";			
RL	Endocrinology 137:4285-4297(1996).			
CC	-!- FUNCTION: THIS IS A SPECIFIC RECEPTOR FOR PARATHYROID HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL CYCLASE. PTH2 MAY BE RESPONSIBLE FOR PTH EFFECTS IN A NUMBER OF PHYSIOLOGICAL SYSTEMS. IT MAY PLAY A SIGNIFICANT ROLE IN PANCREATIC FUNCTION. PTH2 PRESENCE IN NEURONS INDICATES THAT IT MAY FUNCTION AS A NEUROTRANSMITTER RECEPTOR.			
CC	-!- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN BRAIN, ARTERIAL AND CARDIAC ENDOTHELIUM. FOUND AS WELL IN SPERM, IN THE HEAD OF THE EPIDIDYMS. LOWER EXPRESSION IS FOUND IN VASCULAR SMOOTH MUSCLE, EXOCRINE PANCREAS, TESTIS AND PLACENTA.			
CC	-!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U55836; AAC52849.1; -			
DR	GENDB; GCR_1413; -			
DR	INTERPRO; IPR000832; -			
DR	PFAM; PF00002; 7tm_2; 1.			
DR	PRINTS; PR00249; GPCRSECRETIN.			
DR	PROSITE; PS00649; G.PROTEIN.RECEP_F2.1; 1.			
DR	PROSITE; PS00650; G.PROTEIN.RECEP_F2.2; FALSE_NEG.			
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.			
FT	SIGNAL 1 24			
FT	CHAIN 1 24			
FT	DOMAIN 25 546			
FT	DOMAIN 27 143			
FT	TRANSMEM 144 167			
FT	DOMAIN 168 174			
FT	TRANSMEM 175 194			
FT	DOMAIN 195 235			
FT	TRANSMEM 236 273			
FT	DOMAIN 259 273			

34	37	58.7	162	1 CALM_CHLRE	P04352 chlamydomon
35	37	58.7	169	1 CAFL1_DICDI	P54670 dictyosteli
36	37	58.7	218	1 CB25_TETH	P09226 tetrahymena
37	37	58.7	255	1 LACR_LACLA	P18816 lactococcus
38	37	58.7	257	1 CGIC_ORYSA	P33411 oryza sativ
39	37	58.7	522	1 SR54_SCHPO	P21565 schizosacch
40	37	58.7	532	1 CDPK_DAUCA	P28582 daucus caro
41	37	58.7	538	1 CH60_STAEP	P48218 staphylococ
42	37	58.7	543	1 CH60_BACSU	P28598 bacillus su
43	37	58.7	544	1 CH60_BACHD	O50305 bacillus ha
44	37	58.7	1125	1 Y079_HUMAN	P53992 homo sapien
45	36	57.1	54	1 IOVO_CARME	P52242 carpodacus

ALIGNMENTS

THIS PAGE BLANK (USPTO)

C:Keywords: calcium binding; EF hand
F:9-41/Domain: calmodulin repeat homology <EF1>
F:45-77/Domain: calmodulin repeat homology <EF2>
F:82-114/Domain: calmodulin repeat homology <EF3>
F:118-150/Domain: calmodulin repeat homology <EF4>

Query Match	61.9%	Score 39;	DB 2;	Length 151;
Best Local Similarity	70.0%	Pred. No. 14;		
Matches	7;	Conservative	1;	Mismatches
			2;	Indels
				Gaps
				0;

Qy	4	DSGGTITIEE	13
Db	24	DGDGCITVEE	33

Search completed: November 8, 2000, 08:53:21
Job time: 352 sec

RESULT 11
S32106
groEL protein - Lactococcus lactis
C;Species: Lactococcus lactis
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-Aug-1999
C;Accession: S32106
R;Batt, C.A.
submitted to the EMBL Data Library, March 1993
A;Reference number: S32105
A;Accession: S32106
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-542 <BAT>
A;Cross-references: EMBL:X71132; NID:g287869; PIDN:CAA50446.1; PID:g287871
C;Superfamily: chaperonin groEL

Query Match 63.5%; Score 40; DB 2; Length 542;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 QLSDSGTITIEE 13
Db 166 RVGSDGVITIEE 177
::: ||| |||||

RESULT 12
S70013
chaperonin-like protein groEL2 - Synechococcus sp.
C;Species: Synechococcus sp.
C;Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 20-Jun-2000
C;Accession: S70013; S70022
R;Furukii, M.; Tanaka, N.; Hiyama, T.; Nakamoto, H.
Biochim. Biophys. Acta 1294, 106-110, 1996
A;Title: Cloning, characterization and functional analysis of groEL-like gene from thermophilic Synechococcus sp.
A;Reference number: S70013; MUID:96248387
A;Accession: S70013
A;Molecule type: DNA
A;Residues: 1-543 <FUR>
A;Cross-references: GB:D86384; EMBL:DL7354; NID:g1408522; PIDN:BAAL3082.1; PID:g1228069
A;Note: the source is designated as Synechococcus vulcanus
A;Accession: S70022
A;Molecule type: protein
A;Residues: 2-11 <FUR>
A;Note: the source is designated as Synechococcus vulcanus
C;Genetics:
A;Gene: groEL2
C;Superfamily: chaperonin groEL
C;Keywords: heat shock

Query Match 63.5%; Score 40; DB 2; Length 543;
Best Local Similarity 61.5%; Pred. No. 36;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 AOLDSGDTITIEE 13
Db 165 AKVGKGVITIEE 177
::: ||| |||||

RESULT 13
T04252
probable phosphatidylserine decarboxylase (EC 4.1.1.65) F20B18.80 precursor - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 24-Sep-1999
C;Accession: T04252
R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.X.
submitted to the Protein Sequence Database, March 1999
A;Reference number: Z15263
A;Accession: T04252
A;Molecule type: DNA
A;Residues: 1-628 <BEV>

A;Cross-references: EMBL:AL049483
A;Experimental source: cultivar Columbia; BAC clone F20B18
C;Genetics:
A;Map position: 4
A;Introns: 62/3; 89/2; 97/3; 120/3; 150/3; 167/3; 195/3; 226/2; 276/2; 304/3; 342/1;
A;Note: F20B18.80
C;Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 63.5%; Score 40; DB 2; Length 628;
Best Local Similarity 53.8%; Pred. No. 42;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 AOLDSGDTITIEE 13
Db 203 ADLNGDGVITIEE 215
::: ||| |||||

RESULT 14
T08585
calmodulin - soybean
C;Species: Glycine max (soybean)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jan-2000
C;Accession: T08585
R;Lee, S.H.; Kim, J.C.; Choi, Y.J.; Bahk, J.D.; Hong, J.C.; Cho, M.J.
unpublished results 1993, cited by EMBL
A;Reference number: Z16443
A;Accession: T08585
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-150 <LEE>
A;Cross-references: EMBL:L01433; NID:g170075; PID:g170076
C;Genetics:
A;Gene: SCAM-4
C;Superfamily: calmodulin; calmodulin repeat homology
C;Keywords: EF hand
F;8-40/Domain: calmodulin repeat homology <EFH>

Query Match 61.9%; Score 39; DB 2; Length 150;
Best Local Similarity 70.0%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 DSDGTVITIEE 13
Db 23 DGDGTVITIEE 32
::: ||| |||||

RESULT 15
A71409
calmodulin DL3360W - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000
C;Accession: A71409
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D.
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terryn, N.; G.
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomen
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.;
C.; Chalwatzis, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t
A;Reference number: A71400; MUID:98121113
A;Accession: A71409
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-151 <BEV>
A;Cross-references: GB:Z97336; NID:g2244788; PIDN:CAB10243.1; PID:g2244820
C;Genetics:
A;Map position: 4COP9-4G3845
A;Note: DL3360W
C;Superfamily: calmodulin; calmodulin repeat homology

R.Smith, D.R.: Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N. J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functional genome

A:Reference number: A69000; MUID:98037514

A:Accession: H69049

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1044 <MTH>

A:Cross-references: GB:AE000666; NID:g2622478; PIDN:AAB85852.1; PID:g262248

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH1375

C:Superfamily: isoleucine--trna ligase

C:Keywords: aminoacyl-trna synthetase; ligase; protein biosynthesis

Query Match 65.1%; Score 41; DB 2; Length 1044;
Best Local Similarity 66.7%; Pred. No. 48;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AQLSDSGTITIE 12
I I I I I I I I
Db 903 AALSDSGEFTVE 914

RESULT 7
T47245
calcineurin regulatory chain [Imported] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 26-May-2000
C:Accession: T47245
R:Dieminger, M.; Tropshug, M.
submitted to the EMBL Data Library, April 1997
A:Reference number: Z24421
A:Accession: T47245
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-174 <DTF>
A:Cross-references: EMBL:Y12814; PIDN:CAA73345.1
A:Experimental source: strain 74 A
C:Genetics:
A:Gene: Can B
C:Superfamily: calmodulin; calmodulin repeat homology
C:Keywords: EF hand

Query Match 63.5%; Score 40; DB 2; Length 174;
Best Local Similarity 61.5%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AQLSDSGTITIEE 13
I I I I I I I I
Db 143 ADLQDKGISFEE 155

RESULT 8
C64355
conserved hypothetical protein MJ0443 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 08-Oct-1999
C:Accession: C64355
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999
A:Accession: C64355
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A:Residues: 1-227 <BUL>
A:Cross-references: GB:U67495; GB:L77117; NID:g1591134; PIDN:AAB98430.1; PID:g1591147
C:Genetics:
A:Map position: REV397693-397010
C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0443

Query Match 63.5%; Score 40; DB 2; Length 227;
Best Local Similarity 60.0%; Pred. No. 15;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QLSDSGTITI 11
:::||||:|
Db 49 EIDADGTVTI 58

RESULT 9
T28036
hypothetical protein ZK849.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T28036
R:Kershaw, J.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z20460
A:Accession: T28036
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-390 <WIL>
A:Cross-references: EMBL:Z82095; PIDN:CAB05026.1; GSPDB:GN00019; CESP:ZK849.2
A:Experimental source: clone ZK849
C:Genetics:
A:Gene: CESP:ZK849.2
A:Map position: 1
A:Introns: 33/3; 66/3; 103/3; 210/3; 255/2; 301/3; 356/3

Query Match 63.5%; Score 40; DB 2; Length 390;
Best Local Similarity 77.8%; Pred. No. 26;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 DSDGTITIE 12
I I I I I I I I
Db 376 DDDGTVTIE 384

RESULT 10
JN0661
heat shock protein groEL - Lactococcus lactis subsp. lactis
C:Species: Lactococcus lactis subsp. lactis
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 17-Mar-1999
C:Accession: JN0661
R:Kim, S.G.; Batt, C.A.
Gene 127, 121-126, 1993
A:Title: Cloning and sequencing of Lactococcus lactis subsp. lactis groESL operon.
A:Reference number: JN0660; MUID:93252268
A:Accession: JN0661
A:Molecule type: DNA
A:Residues: 1-542 <KIM>
C:Genetics:
A:Gene: groEL
C:Superfamily: chaperonin groEL
C:Keywords: heat shock; molecular chaperone; stress-induced protein

Query Match 63.5%; Score 40; DB 2; Length 542;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QLSDSGTITIEE 13
:: ||| |||||
Db 166 RVSGDGVITIEE 177

Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LQSDGTITIEE 13
:|||||:
Db 426 VNSDGTITIKE 436

RESULT 3

T19643 hypothetical protein C32H11.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T19643; T19646

R:Barlow, K.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19157

A:Accession: T19643

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-195 <WIL>

A:Cross-references: EMBL:282260; PIDN:CAB05134.1; GSPDB:GN00022; CESP:C32H11.8

A:Experimental source: clone C32H11

A:Accession: T19646

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-195 <WIL>

A:Cross-references: EMBL:282260; PIDN:CAB05137.1; GSPDB:GN00022; CESP:C32H11.11

A:Experimental source: clone C32H11

C:Genetics:

A:Gene: CESP:C32H11.8; CESP:C32H11.11

A:Map position: 4

A:Introns: 74/3

Query Match 65.1%; Score 41; DB 2; Length 195;

Best Local Similarity 66.7%; Pred. No. 8.5;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 QLSDGTITIEE 13

:|||||:

Db 48 QLSDGTITIAE 59

RESULT 4

H70737 probable o-sialoglycoprotein endopeptidase - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C:Accession: H70737

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987

A:Accession: H70737

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-344 <COL>

A:Cross-references: GB:277165; GB:AL123456; NID:g3261609; PIDN:CAB01004.1; PID:g1449368

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: gcp

A:Superfamily: O-sialoglycoprotein endopeptidase

Query Match 65.1%; Score 41; DB 2; Length 344;

Best Local Similarity 63.6%; Pred. No. 15;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQLSDGTITTI 11

:||| |||:

Db 1 AQLSDGTITTI 11

Db 19 ARLPDGTVTL 29

RESULT 5

A57429

aldehyde oxidase (EC 1.2.3.1) - Desulfovibrio gigas

C:Species: Desulfovibrio gigas

C:Date: 30-Nov-1995 #sequence_revision 07-Feb-1997 #text_change 11-Jun-1999

C:Accession: A57429; S42565; S35043; S41442

R:Ramao, M.J.; Archer, M.; Moura, I.; Moura, J.J.G.; LeGall, J.; Engh, R.; Schneider,

Science 270, 1170-1176, 1995

A:Title: Crystal structure of the xanthine oxidase-related aldehyde oxidoreductase f

A:Reference number: A57429; MUID:96072968

A:Contents: corrections; annotation; X-ray crystallography, 2.25 angstroms

A:Accession: A57429

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-907 <RAM>

A:Cross-references: EMBL:X7722; NID:g853816; PIDN:CAA54439.1; PID:g853817

A:Note: X-ray structure structure indicated two frame shift errors that were confirmed

R:Thoenes, U.; Flores, O.L.; Neves, A.; Devreese, B.; van Beeumen, J.J.; Huber, R.; R

Eur. J. Biochem. 220, 901-910, 1994

A:Title: Molecular cloning and sequence analysis of the gene of the molybdenum-contain

dehydrogenase.

A:Reference number: S42565; MUID:94192682

A:Accession: S42565

A:Molecule type: DNA; protein

A:Residues: 1-624; AWTALTTPKPGSPSMTAPSPCIRPKTKMARARTSAAWA', 'RRTKPCVPWAWLKRKSSSGPTTP

A:Cross-references: EMBL:X7722; NID:g853816

A:Experimental source: strain ATCC 27774

R:Ramao, M.J.; Barata, B.A.S.; Archer, M.; Lobeck, K.; Moura, I.; Carrondo, M.A.; LeG

Eur. J. Biochem. 215, 729-732, 1993

A:Title: Subunit composition, crystallization and preliminary crystallographic studie

A:Reference number: S35043; MUID:93358896

A:Accession: S35043

A:Molecule type: protein

A:Residues: 1-29, DXX, 33-34 <ROM>

R:Ramao, M.J.; Archer, M.; Moura, I.; Moura, J.J.G.; LeGall, J.; Engh, R.; Schneider,

submitted to the Brookhaven Protein Data Bank, September 1995

A:Reference number: A65095; PDB:1ALO

A:Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 1-699, 'E', 701-

C:Comment: This enzyme contains two 2Fe-2S clusters and a non-covalently bound molybd

C:Genetics:

C:Gene: MOP

C:Complex: homodimer

C:Function:

A:Description: catalyzes the oxidation of a broad range of aldehydes to carboxylic ac

C:Superfamily: Desulfovibrio gigas aldehyde oxidase; ferredoxin [2Fe-2S] homology

C:Keywords: 2Fe-2S; FAD; flavoprotein; homodimer; iron-sulfur protein; metalloprotein

F:1-907/Product: aldehyde oxidase #status predicted <MAT>

F:23-61/Domain: ferredoxin [2Fe-2S] homology <FER1>

F:40, 45, 48, 60/Binding site: 2Fe-2S cluster (Cys) (covalent) #status experimental

F:100, 103, 137, 139/Binding site: 2Fe-2S cluster (Cys) (covalent) #status experimental

F:533/Binding site: molybdopterine cytosine dinucleotide (Arg) #status experimental

F:653, 869/Active site: His, Glu #status predicted

Query Match 65.1%; Score 41; DB 1; Length 907;

Best Local Similarity 63.6%; Pred. No. 41;

Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQLSDGTITTI 11

:||| |||:

Db 636 AELNADGTITV 646

RESULT 6

H69049

isoleucine--trRNA ligase (EC 6.1.1.5) - Methanobacterium thermoautotrophicum (strain D

N)Alternate names: isoleucyl-tRNA synthetase

C:Species: Methanobacterium thermoautotrophicum

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 16-Jul-1999

C:Accession: H69049

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:53:19 ; Search time 99.87 Seconds
(without alignments)
8.261 Million cell updates/sec

Title: US-09-236-468A-2_COPY_24_36

Perfect score: 63

Sequence: 1 AQLDSGDTITIEE 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 182106

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_65.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	63	100.0	550	2 A57519	parathyroid hormon
2	42	66.7	2	JN0097	secreted 45K prote
3	41	65.1	195	2 T19643	hypothetical prote
4	41	65.1	344	2 H70737	probable o-sialogl
5	41	65.1	907	1 A57429	aldehyde oxidase (
6	41	65.1	1044	2 H69049	isoleucine--trna l
7	40	63.5	174	2 T47245	calcineurin regula
8	40	63.5	227	2 C64355	conserved hypothet
9	40	63.5	390	2 T28036	hypothetical prote
10	40	63.5	542	2 JN0661	heat shock protein
11	40	63.5	542	2 S32106	groEL protein - La
12	40	63.5	543	2 S70013	chaperonin-like pr
13	40	63.5	628	2 T04252	probable phosphati
14	39	61.9	150	2 T08595	calmodulin - soybe
15	39	61.9	151	2 A71409	calmodulin DL3360W
16	39	61.9	193	2 T15681	hypothetical prote
17	39	61.9	243	2 A70670	hypothetical prote
18	39	61.9	559	2 T09038	NADH dehydrogenase
19	39	61.9	742	1 H75575	oxidoreductase lsi
20	38	60.3	150	2 T06437	calmodulin - garde
21	38	60.3	195	2 T49012	calmodulin-like pr
22	38	60.3	218	2 C70807	probable fadbl8 pr
23	38	60.3	241	2 S62922	probable membrane
24	38	60.3	324	2 T02109	calmodulin-related
25	38	60.3	479	2 T49871	peroxisomal Ca-dep
26	38	60.3	548	2 F70807	probable fadbl9 pr
27	38	60.3	571	2 T00835	calcium-dependent
28	38	60.3	702	2 C71269	probable polyribon
29	38	60.3	801	2 A47744	diacylglycerol kin

30	38	60.3	884	2 E75489	conserved hypothet
31	38	60.3	910	2 T22050	hypothetical prote
32	38	60.3	924	2 T22044	hypothetical prote
33	38	60.3	2168	2 T30171	ninein - mouse
34	37	58.7	69	2 JCL1310	calmodulin-related
35	37	58.7	131	2 A29422	calmodulin-like pr
36	37	58.7	138	2 S02691	calmodulin B - sea
37	37	58.7	141	2 S02690	calmodulin A - sea
38	37	58.7	147	1 MCJZR	calmodulin - sea p
39	37	58.7	148	1 MCRB	calmodulin - rabbi
40	37	58.7	148	1 MCRB	calmodulin - bovin
41	37	58.7	148	1 MCON	calmodulin - salmo
42	37	58.7	148	1 MCLQ	calmodulin - migra
43	37	58.7	148	1 MCSFCU	calmodulin - sea c
44	37	58.7	148	1 MCAZS	calmodulin - sea s
45	37	58.7	148	1 MCSW	calmodulin - scall

ALIGNMENTS

RESULT 1

A57519

parathyroid hormone receptor 2 precursor - human

N:Alternate names: PTH2 receptor

C:Species: Homo sapiens (man)

C>Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 24-Nov-1999

C:Accession: A57519

R:Usdin, T.B.; Gruber, C.; Bonner, T.I.

J. Biol. Chem. 270, 15455-15458, 1995

A:Title: Identification and functional expression of a receptor selectively recognizi

A:Reference number: A57519; MUID:95318121

A:Accession: A57519

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-550 <USD>

A:Cross-references: GB:025128; NID:g887966; PIDN:AAC50157.1; PID:g887967

C:Genetics:

A:Gene: GDB:PTH2; PTHR2

A:Cross-references: GDB:731977; OMIM:601469

A:Map position: 2q33-2q33

C:Superfamily: glucagon receptor

C:Keywords: hormone receptor

Query Match 100.0%; Score 63; DB 2; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQLDSGDTITIEE 13

Db 24 AQLDSGDTITIEE 36

RESULT 2

JN0097

secreted 45K protein precursor - Lactococcus lactis

C:Species: Lactococcus lactis

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999

C:Accession: JN0097

R:van Asseldonk, M.; Rutten, G.; Oteman, M.; Siezen, R.J.; de Vos, W.M.; Simons, G.

Gene 95, 155-160, 1990

A:Title: Cloning of usp45, a gene encoding a secreted protein from Lactococcus lactis

A:Reference number: JN0097; MUID:91071599

A:Accession: JN0097

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-461 <VAN>

A:Cross-references: GB:M35374

Query Match 66.7%; Score 42; DB 2; Length 461;
Best Local Similarity 72.7%; Pred. No. 14;

Query Match 58.7%; Score 37; DB 2; Length 642;
Best Local Similarity 70.0%; Pred. No. 64;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 DSDGTTITIEE 13
| | | | | : |
Db 252 DGDGTTITKE 261

Search completed: November 8, 2000, 08:49:15
Job time: 108 sec

;
; TITLE OF INVENTION: DISEASE ASSOCIATED CALMODULIN PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/963,409
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0418 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 825635
; US-08-963-409-5

Query Match 58.7%; Score 37; DB 3; Length 152;
Best Local Similarity 70.0%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 DSDGTITIEE 13
| | | | | : |
Db 23 DGDGTITTKK 32

RESULT 14
US-08-253-2
; Sequence 2, Application US/08818253
; Patent No. 5998204
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: DETECTION OF ANALYTES
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,253
; FILING DATE: 14-MAR-1997

;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/043001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 642 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-818-253-2

Query Match 58.7%; Score 37; DB 2; Length 642;
Best Local Similarity 70.0%; Pred. No. 64;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 DSDGTITIEE 13
| | | | | : |
Db 252 DGDGTITTKK 261

RESULT 15
US-08-818-253-6
; Sequence 6, Application US/08818253
; Patent No. 5998204
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: DETECTION OF ANALYTES
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,253
; FILING DATE: 14-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/043001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 642 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-818-253-6

; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 149 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 385234
US-08-963-409-3

Query Match 58.7%; Score 37; DB 3; Length 149;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 DSDGTITIEE 13
| | | | | : |
Db 23 DGDGTTTKE 32

RESULT 11
US-08-963-409-4
; Sequence 4, Application US/08963409
; Patent No. 6046315
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: DISEASE ASSOCIATED CALMODULIN PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/963,409
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0418 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 149 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1334203
US-08-963-409-4

Query Match 58.7%; Score 37; DB 3; Length 149;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 DSDGTITIEE 13
| | | | | : |
Db 23 DGDGTTTKE 32

RESULT 12
US-08-641-873-20
; Sequence 20, Application US/08641873
; Patent No. 6117976
; GENERAL INFORMATION:
; APPLICANT: Neri, D.
; APPLICANT: Winter, GP.
; APPLICANT: de Lallia, C.
; TITLE OF INVENTION: Manufacture and use of polypeptides tagged
; with binding molecules
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/641,873
; FILING DATE: 02-MAY-1996
; CLASSIFICATION: 435
; GOIN 33/68
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB94/02420
; FILING DATE: 04-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9322772.6
; FILING DATE: 04-NOV-1993
; APPLICATION DATA:
; APPLICATION NUMBER: GB 9405927.6
; FILING DATE: 25-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: David W. Clough
; REGISTRATION NUMBER: 36,107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 149 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-641-873-20

Query Match 58.7%; Score 37; DB 3; Length 149;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 DSDGTITIEE 13
| | | | | : |
Db 23 DGDGTTTKE 32

RESULT 13
US-08-963-409-5
; Sequence 5, Application US/08963409
; Patent No. 6046315
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.

```
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/459,595
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1577/CIP/DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 142 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..142
; OTHER INFORMATION: /note= "human calmodulin protein
; OTHER INFORMATION: sequence as shown in Figure 33."
;
US-08-459-504B-24

Query Match 58.7% Score 37; DB 3; Length 142;
Best Local Similarity 70.0% Pred. No. 11;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 DSDGTITTEE 13
Db 19 DGDGTITTKE 28

RESULT 9
US-08-459-444-24
; Sequence 24, Application US/08459444A
; Patent No. 6121014
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; Desai, Nalini M.
; Lewis, Kelly S.
; Kramer, Vance C.
; Warren, Gregory W.
; Evola, Stephen V.
; Crossland, Lyle D.
; Wright, Martha S.
; Merlin, Ellis J.
; Launis, Karen L.
; TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED
; NUCLEIC ACID CODING SEQUENCE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6121014artis Agribusiness Biotechnology Research, Inc.
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,444A
; FILING DATE: 02-Jun-1995
```

```
;
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-18805/Pl/CGC1577/CIP/DIV6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 142 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEtical: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..142
; OTHER INFORMATION: /note= "human calmodulin protein
; OTHER INFORMATION: sequence as shown in Figure 33."
;
US-08-459-444-24

Query Match 58.7% Score 37; DB 3; Length 142;
Best Local Similarity 70.0% Pred. No. 11;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 DSDGTITTEE 13
Db 19 DGDGTITTKE 28

RESULT 10
US-08-963-409-3
; Sequence 3, Application US/08963409
; Patent No. 6046315
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: DISEASE ASSOCIATED CALMODULIN PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/963,409
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0418 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
```


TELEPHONE: (919)541-8582
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 142 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Protein
LOCATION: 1..142
OTHER INFORMATION: /note= "human calmodulin protein
US-08-459-448A-24

Query Match 58.7%; Score 37; DB 2; Length 142;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 0;

QY 4 DSDGTITIEE 13
| | | | | : |
Db 19 DGDGTITTKK 28

RESULT 7
US-08-459-595A-24
Sequence 24, Application US/08459595A
Patent No. 6018104

GENERAL INFORMATION:
APPLICANT: Koziel, Michael G.
APPLICANT: Desai, Nalini M.
APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.
APPLICANT: Evola, Stephen V.
APPLICANT: Crossland, Lyle D.
APPLICANT: Wright, Martha S.
APPLICANT: Merlin, Ellis J.
APPLICANT: Launis, Karen L.
APPLICANT: Rothstein, Steven J.
APPLICANT: Bowman, Cindy G.
APPLICANT: Dawson, John L.
APPLICANT: Dunder, Erik M.
APPLICANT: Pace, Gary M.
APPLICANT: Suttie, Janet L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6018104artis Corporation
STREET: Patent & Trademark Dept., 520 White Plains
STREET: Rd., POB 2005
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591-9005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,595A
FILING DATE: 02-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/772,027

FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: 40403
REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8582
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 142 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Protein
LOCATION: 1..142
OTHER INFORMATION: /note= "human calmodulin protein
US-08-459-595A-24

Query Match 58.7%; Score 37; DB 3; Length 142;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 0;

QY 4 DSDGTITIEE 13
| | | | | : |
Db 19 DGDGTITTKK 28

RESULT 8
US-08-459-504B-24
Sequence 24, Application US/08459504B
Patent No. 6075185
GENERAL INFORMATION:
APPLICANT: Koziel, Michael G.
APPLICANT: Desai, Nalini M.
APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.
APPLICANT: Evola, Stephen V.
APPLICANT: Crossland, Lyle D.
APPLICANT: Wright, Martha S.
APPLICANT: Merlin, Ellis J.
APPLICANT: Launis, Karen L.
APPLICANT: Rothstein, Steven J.
APPLICANT: Bowman, Cindy G.
APPLICANT: Dawson, John L.
APPLICANT: Dunder, Erik M.
APPLICANT: Pace, Gary M.
APPLICANT: Suttie, Janet L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6075185artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,504B
FILING DATE:
CLASSIFICATION:

; TELEFAX: (914)347-5769
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-186-222-2

Query Match 66.7%; Score 42; DB 1; Length 461;
Best Local Similarity 72.7%; Pred. No. 5.5;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LDSGTTITIEE 13
Db 426 VNSDGTITIKE 436
:::|||||||:|

RESULT 5

US-07-951-715A-24
; Sequence 24, Application US/07951715A
; Patent No. 5625136

; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.

; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE

; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CIBA-CEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/951,715A
; FILING DATE: 25-SEP-1992
; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943

; REFERENCE/DOCKET NUMBER: S-18805/A/CGC 1577/CIP
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (919)541-8615

; TELEFAX: (919)541-8689

; INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 142 amino acids

; TYPE: amino acid

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..142
; OTHER INFORMATION: /note= "human calmodulin protein
; sequence as shown in Figure 33."

US-07-951-715A-24

Query Match 58.7%; Score 37; DB 1; Length 142;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 4 DSDGTTITIEE 13
Db 19 DGDGTTITKE 28
| ||||| :|

RESULT 6

US-08-459-448A-24

; Sequence 24, Application US/08459448A
; Patent No. 5859336

; GENERAL INFORMATION:

; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.

; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE

; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NO. 5859336artis Corporation
; STREET: Patent & Trademark Dept., 520 White Plains
; STREET: Rd., POB 2005
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-9005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/459,448A

; FILING DATE: 02-JUN-1995

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/951,715

; FILING DATE: 25-SEP-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/772,027

; FILING DATE: 04-OCT-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Pace, Gary M.

; REGISTRATION NUMBER: 40403

; REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV4

; TELECOMMUNICATION INFORMATION:

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:49:14 ; Search time 97.15 seconds
(without alignments)
2.243 Million cell updates/sec

Title: us-09-236-468a-2_copy_24_36

Perfect score: 63

Sequence: 1 AQLSDSGTITIEE 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 164575

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	63	100.0	37	3	US-08-468-011A-25
2	63	100.0	541	3	US-08-468-011A-2
3	63	100.0	541	4	PCT-US95-07085-2
4	42	66.7	461	1	US-08-186-222-2
5	37	58.7	142	1	US-07-951-715A-24
6	37	58.7	142	2	US-08-459-448A-24
7	37	58.7	142	3	US-08-459-595A-24
8	37	58.7	142	3	US-08-459-504B-24
9	37	58.7	142	3	US-08-459-444-24
10	37	58.7	149	3	US-08-963-409-3
11	37	58.7	149	3	US-08-963-409-4
12	37	58.7	149	3	US-08-641-873-20
13	37	58.7	152	3	US-08-963-409-5
14	37	58.7	642	2	US-08-818-253-2
15	37	58.7	642	2	US-08-818-253-6
16	37	58.7	652	2	US-08-818-253-4
17	37	58.7	656	2	US-08-818-253-8
18	36	57.1	191	3	US-08-764-563-1
19	36	57.1	333	1	US-08-436-463-6
20	36	57.1	333	1	US-08-024-253-6
21	36	57.1	408	1	US-07-951-715A-21
22	36	57.1	408	2	US-08-459-448A-21
23	36	57.1	408	3	US-08-459-595A-21
24	36	57.1	408	3	US-08-459-504B-21
25	36	57.1	408	3	US-08-459-444-21
26	36	57.1	464	1	US-07-951-715A-22
27	36	57.1	464	2	US-08-459-448A-22
28	36	57.1	464	3	US-08-459-595A-22

Sequence 22, Appl
Sequence 0, Appl
Sequence 2, Appl
Sequence 1, Appl
Sequence 2, Appl
Sequence 30, Appl
Sequence 6, Appl
Sequence 6, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 10, Appl
Sequence 10, Appl
Sequence 5, Appl
Sequence 4, Appl
Sequence 8, Appl
Sequence 23, Appl

ALIGNMENTS

```

RESULT 1
US-08-468-011A-25
; Sequence 25, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468.011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-011A-25

```

Query Match 100.0%; Score 63; DB 3; Length 37;
Best Local Similarity 100.0%; Pred. No. 4.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQLSDSGTITIEE 13
|||||

SQ Sequence 151 AA;

Query Match 58.7%; Score 37; DB 9; Length 151;
 Best Local Similarity 70.0%; Pred. No. 16;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 DSDGTITIEE 13
 | | | | | : |
 Db 25 dgdgtittke 34

RESULT 12

R78526
 ID R78526 standard; Protein; 375 AA.

AC R78526;

DT 31-JAN-1996 (first entry)

DE GST-calmodulin fusion protein.

KW Human; calmodulin; fibrosarcoma cell; primer; PCR; amplification;
 maltose; protein A; glutathione-S-transferase; fusion protein; activity;
 phosphodiesterase activating method; binding protein; inhibitor.

OS Synthetic.

XX Key Location/Qualifiers
 FH Peptide 228..375
 FT /note= "calmodulin peptide"

PN JP07126297-A.

XX 16-MAY-1995.

XX 05-NOV-1993; 93JP-0299041.

XX 05-NOV-1993; 93JP-0299041.

XX (SAGA) SAGAMI CHEM RES CENTRE.

XX WPI; 1995-212960/28.

XX N-PSDB; Q95552.

PT Fusion of calmodulin (CAM) with specific binding peptide - has same
 activity as CAM, useful for screening for CAM inhibitors, measuring
 CAM binding protein activity, etc.

PS Examples; Page 10-11; 12pp; Japanese.

XX The amino acid sequence of the glutathione-S-transferase-human calmodulin
 fusion protein from clone pGSTRCAM. The clone was generated by cloning
 the calmodulin cDNA (Q95549) into the plasmid pGEX-5X-1. The calmodulin
 cDNA was also cloned into the plasmids pRIT2T and pMAL(RTM)-2 to generate
 pMALCAM (Q95551) and pMALCAM (Q95550) which produce protein A-
 (R78525) and maltose binding protein-calmodulin (R78526) fusion proteins,
 respectively. The fusion proteins produced retain the same activity as
 calmodulin as measured by the phosphodiesterase activating method. The
 fusion protein, obtainable at high purity, can be used to isolate
 calmodulin binding proteins, measure their activity and screen for
 calmodulin inhibitors.

SQ Sequence 375 AA;

Query Match 58.7%; Score 37; DB 16; Length 375;
 Best Local Similarity 70.0%; Pred. No. 49;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 DSDGTITIEE 13
 | | | | | : |
 Db 249 dgdgtittke 258

RESULT 13

R78525
 ID R78525 standard; Protein; 409 AA.

XX R78525;

DT 31-JAN-1996 (first entry)

DE Protein A-calmodulin fusion protein.

XX Human; calmodulin; fibrosarcoma cell; primer; PCR; amplification;
 maltose; protein A; glutathione-S-transferase; fusion protein; activity;
 phosphodiesterase activating method; binding protein; inhibitor.

OS Synthetic.

XX Key Location/Qualifiers
 FH Peptide 262..409
 FT /note= "calmodulin peptide"

PN JP07126297-A.

XX 16-MAY-1995.

XX 05-NOV-1993; 93JP-0299041.

XX 05-NOV-1993; 93JP-0299041.

XX (SAGA) SAGAMI CHEM RES CENTRE.

XX WPI; 1995-212960/28.

XX N-PSDB; Q95551.

PT Fusion of calmodulin (CAM) with specific binding peptide - has same
 activity as CAM, useful for screening for CAM inhibitors, measuring
 CAM binding protein activity, etc.

PS Examples; Page 8-10; 12pp; Japanese.

XX The amino acid sequence of the protein A-human calmodulin fusion protein
 from clone pPACAM. The clone was generated by cloning the calmodulin
 cDNA (Q95549) into the plasmid pRIT2T. The calmodulin cDNA was also
 cloned into the plasmids pMAL(RTM)-2 and pGEX-5X-1 to generate plasmids
 pMALCAM (Q95550) and pGSTRCAM (Q95552) which produce maltose binding
 protein- (R78524) and glutathione-S-transferase-calmodulin (R78526)
 fusion proteins, respectively. The fusion proteins produced retain the
 same activity as calmodulin as measured by the phosphodiesterase
 activating method. The fusion protein, obtainable at high purity, can be
 used to isolate calmodulin binding proteins, measure their activity and
 screen for calmodulin inhibitors.

SQ Sequence 409 AA;

Query Match 58.7%; Score 37; DB 16; Length 409;
 Best Local Similarity 70.0%; Pred. No. 54;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 DSDGTITIEE 13
 | | | | | : |
 Db 283 dgdgtittke 292

RESULT 14

W92993
 ID W92993 standard; Protein; 416 AA.

XX W92993;

DT 18-MAY-1999 (first entry)

XX

```

XX DE Human calmodulin protein.
XX KW Human; calmodulin; fibrosarcoma cell; primer: PCR; amplification;
XX KW maltose; protein A; glutathione-S-transferase; fusion protein; activity;
XX KW phosphodiesterase activating method; binding protein; inhibitor.
XX OS Homo sapiens.
XX PN JP07126297-A.
XX PD 16-MAY-1995.
XX PF 05-NOV-1993; 93JP-0299041.
XX PR 05-NOV-1993; 93JP-0299041.
XX PA (SAGA ) SAGAMI CHEM RES CENTRE.
XX DR WPI; 1995-212960/28.
XX DR N-PSDB; Q95549.
XX PT Fusion of calmodulin (CAM) with specific binding peptide - has same
XX PT activity as CAM, useful for screening for CAM inhibitors, measuring
XX PT CAM binding protein activity, etc.
XX PS Examples; Page 5-6; 12pp; Japanese.
XX CC The amino acid sequence of the human calmodulin cDNA clone pKAL-CAM.
XX CC The clone was isolated from the human fibrosarcoma cell line HT-1080
XX CC using primers Q95533-4 to PCR amplify the coding region. The coding
XX CC region of the gene was inserted into the plasmids pMAL(RTM)-2, pRIT2T
XX CC and pGEX-5X-1 to generate plasmids pMALCAM (Q95550), pPACAM (Q95551)
XX CC and pGSTCAM (Q95552) respectively. These plasmids express maltose-
XX CC (R78524), protein A- (R78525) and glutathione-S-transferase-calmodulin
XX CC (R78526) fusion proteins, respectively. The fusion proteins produced
XX CC retain the same activity as calmodulin as measured by the
XX CC phosphodiesterase activating method. The fusion protein, which can be
XX CC obtained at high purity, can be used to isolate calmodulin binding
XX CC proteins, measure their activity and screen for calmodulin inhibitors.
XX SQ Sequence 149 AA;

Query Match 58.7%; Score 37; DB 16; Length 149;
Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 DSDGTITIEE 13
DB 23 dgdgtittke 32

RESULT 10
Y50797
ID Y50797 standard; protein; 149 AA.
XX AC Y50797;
XX DT 17-FEB-2000 (first entry)
XX DE Human calmodulin/NEF protein binding inhibitor 2.
XX KW NEF protein; calmodulin; negative factor protein; binding inhibitor;
XX KW diagnosis; detection; infection; treatment; HIV.
XX OS Homo sapiens.
XX PN WO9957136-A2.
XX PD 11-NOV-1999.
XX PF 06-MAY-1999; 99WO-EP03105.

```

```

XX PR 06-MAY-1998; 98DE-1020224.
XX PA (SCHO/) SCHOTT M.
XX PA (SCHO/) SCHORR J.
XX PA (ANTZ/) ANTZ C.
XX PI Schott M, Schorr J, Antz C;
XX DR WPI; 2000-038789/03.
XX PT Binding agents used for treatment, prevention and diagnosis of human
XX PT immune deficiency virus infection -
XX PS Disclosure; Fig 3; 29pp; German.
XX CC This invention describes novel binding partners (A) (i) for negative
XX CC factor protein (Nef) that competitively inhibit binding of calmodulin (I)
XX CC to Nef, or (ii) for (I) that competitively inhibit binding of Nef to (I).
XX CC (A) are used for prevention, diagnosis (by specific detection of intra-
XX CC and/or extra-cellular Nef, including staging of infection), and/or
XX CC treatment of infections with human immune deficiency virus (HIV).
XX SQ Sequence 149 AA;

Query Match 58.7%; Score 37; DB 21; Length 149;
Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 DSDGTITIEE 13
DB 23 dgdgtittke 32

RESULT 11
P80162
ID P80162 standard; protein; 151 AA.
XX AC P80162;
XX DT 14-NOV-1990 (first entry)
XX DE Biosynthetic multifunctional protein.
XX KW Biosynthetic multifunctional protein; biosynthetic antibody binding site;
XX KW protein trailer; calmodulin.
XX PN WO8809344-A.
XX PD 01-DEC-1988.
XX PF 19-MAY-1988; 88WO-US01737.
XX PR 21-MAY-1987; 87US-0052800.
XX PA (CREA-) CREATIVE BIOMOLECULES INC.
XX PI Huston JS, Oppermann H;
XX DR WPI; 1988-353928/49.
XX DR N-PSDB; N80188.
XX PT Recombinant multifunctional protein - having antibody binding site and a
XX PT sequence for biological activity, ion sequestering or binding to a
XX PT solid support.
XX PS Disclosure; ; 115pp; English.
XX CC The sequence is a biosynthetic multifunctional protein including a
XX CC biosynthetic antibody binding site and calmodulin protein trailer linked
XX CC via a spacer sequence.

```

FT /label= KR6
FT /note= "Beta-ketoreductase domain: part of extender
FT module 6"
FT 3188..3269
FT /label= ACP6
FT /note= "Acyl carrier protein domain: part of extender
FT module 6"
FT 3291..3713
FT /label= KS7
FT /note= "Beta-ketosynthase domain: part of extender
FT module 7"
FT 3825..4153
FT /label= A17
FT /note= "Acyl transferase domain: part of extender
FT module 7"
FT 4344..4638
FT /label= KR7
FT /note= "Beta-ketoreductase domain: part of extender
FT module 7"
FT 4725..4806
FT /label= ACP7
FT /note= "Acyl carrier protein domain: part of extender
FT module 7"
XX
PN WO9946387-A1.
XX
XX
PD 16-SEP-1999.
XX
PF 16-FEB-1999; 99WO-US03212.
XX
PR 09-MAR-1998; 98US-0036987.
XX
XX (DOWC) DOW AGROSCIENCES LLC.
XX
PI Baltz RH, Broughton MC, Crawford KP, Madduri K, Merlo DJ;
PI Treadway PJ, Turner JR, Waldron C;
XX
DR WPI; 1999-551414/46.
DR N-PSDB; 221501.
XX
XX
FT New spinosyn biosynthetic genes from Saccharopolyspora spinosa, useful
FT for production of insecticidal spinosyn compounds -
PS Claim 1; Page 113-129; 190pp; English.
XX
XX This is the amino acid sequence of the product of the spnD gene. The
CC protein is involved in spinosyn biosynthesis. The SpnD gene is one of 23
CC genes and open reading frames contained in an 80kb DNA sequence 221501.
CC Spinosyns are insecticidal microbicides which are useful for the control of
CC arachnids, nematodes and insects. Biosynthesis of spinosyns occurs via
CC stepwise condensation and modification of carboxylic acid precursors
CC generating a linear polyketide which is modified further. The DNA
CC sequence contains a central region of approximately 55kb which has
CC homology to the DNA encoding the polyketide synthases (PKS) of known
CC macrolide producers. The spinosyn PKS DNA region consists of 5 ORFs with
CC stop codons at the end of acyl carrier protein (ACP) domains. Together
CC the PKS polypeptides (Y39297-Y39301) of which this sequence is one, form
CC a complex consisting of an initiator module, spnA, and several extender
CC modules spnB-spnE. Each extender module adds a specific acetyl Co-A
CC precursor to a growing polyketide chain, and modifies the beta-keto group
CC in a specific manner. A module in a PKS polypeptide consists of several
CC domains with specific functions. The initiator module has an acyl
CC transferase (AT) domain, and an acyl carrier protein (ACP) domain. The
CC extender modules have the same domains plus a beta-ketosynthase (KS)
CC domain and optionally a beta-ketoreductase domain, a dehydratase (DH)
CC domain, and an enoyl reductase (ER) domain. The last extender module
CC terminates with a thiolester domain. The products of the genes present in
CC the upstream region the PKS genes have been assigned names spnF-spnS
CC Y39302-Y39315 and are responsible for different modifications in spinosyn
CC biosynthesis. There are also two ORFs ORF115 and ORF116 present
CC immediately upstream of spnS, producing polypeptides Y39316-Y39317, and
CC two ORFs ORF1 and ORF2 present downstream of the PKS region producing
CC polypeptides Y39318-Y39319. The genes are useful to improve yields of

CC spinosyns, and for creating new spinosyns e.g. by mutagenesis, or
CC interruption of steps in spinosyn biosynthesis. The modified spinosyns
CC may be a new insect control agent or serve as substrates for further
CC chemical modification and the creation of new semi-synthetic spinosyns.
CC The genes are also useful to isolate similar sequences from S. spinosa or
CC other species by hybridization.

XX
SQ Sequence 4928 AA;

Query Match 61.9%; Score 39; DB 20; Length 4928;
Best Local Similarity 70.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 QLSDSGTITI 11
| | | : | | | : |
Db 1416 qltdgtvli 1425

RESULT 8
R04583
ID R04583 standard; protein; 149 AA.
XX
AC R04583;

DT 27-SEP-1990 (first entry)

DE Recombinant calmodulin.

KW Calmodulin.

OS Rattus rattus.

XX JP02092286-A.

PN 03-APR-1990.

XX 30-SEP-1988; 88JP-0246239.

XX 30-SEP-1988; 88JP-0246239.

XX (KIRI) KIRIN BREWERY KK.

DR WPI; 1990-144900/19.

DR N-PSDB; Q04331.

XX Calmodulin prepn. - by culturing transformed E. coli.

XX Disclosure; ; p; Japanese.

CC The recombinant calmodulin is produced by cloning the gene into an
CC expression vector and culturing to produce the protein.
CC See also Q04332-4.

XX
SQ Sequence 149 AA;

Query Match 58.7%; Score 37; DB 11; Length 149;
Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 DSDGTITIEE 13
| | | | | : |
Db 23 dgdgttttke 32

RESULT 9
R78523
ID R78523 standard; Protein; 149 AA.
XX
AC R78523;

DT 31-JAN-1996 (first entry)

XX PS Claim 1; Fig 1; 55pp; English.

XX CC Proteins Y52202-Y52217 encompass a novel family of proteins

XX CC designated the ygd family, after the name given to the Escherichia

XX CC coli family member. These proteins are essential for the survival of

XX CC both Gram negative and Gram positive bacteria, although no function has

XX CC as yet been ascribed to these proteins. The ygd proteins, fragments of

XX CC ygd proteins (for example, fragments encompassing one or more

XX CC conserved ygd motifs such as Y52218-Y52284) and nucleotides encoding

XX CC them can be used to identify antagonists and broad spectrum antibacterial

XX CC compounds. These antagonists and compounds can be used to treat a

XX CC wide range of bacterial infections. New antibiotics are urgently needed,

XX CC as serious bacterial infections and antibiotic resistant strains are

XX CC becoming increasingly prevalent. The proteins of the invention are

XX CC essential proteins for bacterial viability, and represent new targets

XX CC for antibiotics.

XX SQ Sequence 344 AA;

Query Match 65.1%; Score 41; DB 21; Length 344;
 Best Local Similarity 63.6%; Pred. No. 8.1;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AQLDSGTITIT 11
 I: I I I I I:
 Db 19 arldqdgvtvl 29

RESULT 6

W21782 ID W21782 standard; Protein; 243 AA.

XX AC W21782;

DT 11-MAR-1998 (first entry)

XX DE Protein encoded by ORF E of GS region of M. tuberculosis cosmid MTO24.

XX KW GS; pathogenicity island; pathogenic protein; mycobacterial disease;

XX KW cellular recognition receptor; pathogenic mycobacteria; Crohn's disease;

XX KW vaccine; inflammatory disease; sarcoidosis; Johne's disease; ss.

XX OS Mycobacterium tuberculosis.

XX PN WO9723624-A2.

XX PD 03-JUL-1997.

XX PF 23-DEC-1996; 96WO-GB03221.

XX PR 21-DEC-1995; 95GB-0026178.

XX PA (SGEO-) ST GEORGE'S HOSPITAL MEDICAL SCHOOL.

XX PI Doran T, Ford J, Hermon-Taylor J, Loughlin M, Millar D;
 XX PI Sumar N, Tizard M;

XX DR WPI; 1997-351061/32.
 XX DR N-PSDB; T74476.

XX PT New isolated pathogenicity island from mycobacteria - used to
 XX PT develop products for detection, diagnosis, prevention and treatment
 XX PT of mycobacteria infections

XX PS Claim 1; Page 55; 62pp; English.

XX CC The present sequence represents the protein encoded by open reading frame
 XX CC (ORF) E, from M. tuberculosis. This ORF E has been found to have homology
 XX CC with ORF E of a novel polynucleotide sequence designated "GS". GS is a
 XX CC pathogenicity island of 8 kb of DNA comprising a core region of 5.75 kb
 XX CC with multiple ORFs and an adjacent transmissible element of 2.5 kb. The

CC ORFs, and also the transmissible element, encode proteins which may be
 CC linked to pathogenicity, such as providing receptors for cellular
 CC recognition. GS was discovered and characterised using differential DNA
 CC analysis technology. It is found within Mycobacterium paratuberculosis
 CC and it has also been identified in Mycobacterium avium subspecies
 CC silvaticum. These pathogenic mycobacteria cause chronic inflammation of
 CC the intestine and Crohn's disease in humans. The protein products of the
 CC ORFs of GS can be used for detecting mycobacteria or for diagnosing,
 CC treating or preventing mycobacterial disease. In particular they can be
 CC used as vaccines for inflammatory diseases such as Crohn's disease or
 CC sarcoidosis in humans or Johne's disease in animals.

XX SQ Sequence 243 AA;

Query Match 61.9%; Score 39; DB 18; Length 243;
 Best Local Similarity 87.5%; Pred. No. 12;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DSDGTITIT 11
 I: I I I I I:
 Db 98 dsdgtvti 105

RESULT 7

Y39300 ID Y39300 standard; Protein; 4928 AA.

XX AC Y39300;

XX DT 01-DEC-1999 (first entry)

XX DE SpnD a polyketide synthase.

XX KW Spinosyn biosynthetic enzyme; open reading frame; ORF; insecticidal
 XX KW microtides; arachnid; nematode; insect; polyketide; polyketide synthase;
 XX KW PKS; extender module; initiator module; acyl transferase domain; AT;
 XX KW acyl carrier protein; ACP; beta-ketosynthase domain; KS; KR;
 XX KW dehydratase domain; DH; enoyl reductase domain; ER; beta-ketoreductase;
 XX KW insecticide.

XX OS Saccharopolyspora spinosa.

XX FH Key Location/Qualifiers

FT Domain 1..424

FT /label= KS5

FT /note= "Beta-ketosynthase domain: part of extender

FT module 5"

FT Domain 539..866

FT /label= AT5

FT /note= "Acyl transferase domain: part of extender

FT module 5"

FT Domain 893..1078

FT /label= DH5

FT /note= "Dehydratase domain: part of extender

FT module 5"

FT Domain 1384..1565

FT /label= KR5

FT /note= "Beta-ketoreductase domain: part of extender

FT module 5"

FT Domain 1645..1726

FT /label= ACP5

FT /note= "Acyl carrier protein domain: part of extender

FT module 5"

FT Domain 1748..2172

FT /label= KS6

FT /note= "Beta-ketosynthase domain: part of extender

FT module 6"

FT Domain 2283..2613

FT /label= AT6

FT /note= "Acyl transferase domain: part of extender

FT module 6"

FT Domain 2916..3095

hyper-calcaemia, hypo- or hyper-phosphatemia, kidney stones, etc
 Claim 9; Fig 1A-E; 62pp; English.

A novel 7-transmembrane receptor (W12695) has been identified as a human G-protein parathyroid hormone (PTH) receptor, designated HLRPG74. It shows 48.2% homology to the human PTH receptor. Its amino acid sequence was deduced from a cDNA clone (T59619) isolated from a human T cell lymphoma tissue cDNA library. Recombinant HLRPG74 can be produced in transformed host cells and used to screen for (ant)agonist cpds. Agonists may be used to prevent or treat e.g. hypocalcaemia, hyperphosphatemia, hypoparathyroidism and chronic tetany by stimulating an increase in serum calcium levels. Antagonists can be used to inhibit the receptor e.g. for the treatment of osteoporosis, hypercalcaemia, hypoparathyroidism, hypophosphatemia, kidney stone, nephroliasis.

Sequence 541 AA;

Query Match 100.0%; Score 63; DB 18; Length 541;
 Best Local Similarity 100.0%; Pred. No. 0.0013;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AQLDSGDTITIEE 13
 Db 24 aqlsdsgdtitlee 36

RESULT 2
 Y31625
 ID Y31625 standard; Protein; 214 AA.

AC Y31625;

XX 11-OCT-1999 (first entry)

XX Human calcineurin B-like protein.

XX calcineurin B-like protein; CAB; calcium binding protein; CBP; EF-hand;
 KW apoptotic disorder; neoplastic disorder; immunological disorder;
 KW reproductive disorder; acquired immunodeficiency syndrome; AIDS;
 KW Alzheimer's disease; Parkinson's disease; leukaemia; lymphoma;
 KW melanoma; myeloma; Addison's disease; anaemia; infertility;
 KW prolactin production; oestrous cycle disruption.

XX Homo sapiens.

OS Key Location/Qualifiers

XX Domain 123...135

FT /label= EF_hand

FT /note= "calcium binding domain"

FT Modified-site 2

FT /label= myristoylation_site

FT Modified-site 6

FT /label= casein_kinase_II_phosphorylation_site

FT Modified-site 100

FT /label= casein_kinase_II_phosphorylation_site

FT Modified-site 101

FT /label= casein_kinase_II_phosphorylation_site

FT Modified-site 110

FT /label= casein_kinase_II_phosphorylation_site

FT Modified-site 124

FT /label= casein_kinase_II_phosphorylation_site

FT Modified-site 131

FT /label= casein_kinase_II_phosphorylation_site

FT Modified-site 156

FT /label= casein_kinase_II_phosphorylation_site

FT Modified-site 183

FT /label= casein_kinase_II_phosphorylation_site

FT Modified-site 41

FT /label= protein_kinase_C_phosphorylation_site

FT Modified-site 183

FT /label= protein_kinase_C_phosphorylation_site
 FT 183
 FT /label= protein_kinase_C_phosphorylation_site

XX WO9937790-A1.

XX 29-JUL-1999.

XX 19-JAN-1999; 99WO-US01079.

XX 21-JAN-1998; 98US-0010378.

XX (INCY-) INCYTE PHARM INC.

XX Corley NC, Lal P, Tang YT;

XX WPI; 1999-469138/39.

XX N-PSDB; Z08993.

New human calcineurin B-like protein (CAB) useful for diagnosing, preventing and treating apoptotic, neoplastic, immunological and reproductive disorders

Claim 1; Fig 1; 70pp; English.

This sequence represents human calcineurin B-like protein (CAB). Partial cDNA clones were isolated from a bone marrow cDNA library and sequenced. Primers were designed to extend the partial sequence to full length using selected human cDNA libraries. This sequence represents a consensus. The CAB protein is a member of the calcium binding protein (CBP) superfamily, characterised by the presence of an EF-hand calcium binding domain. The CAB protein shares chemical and structural homologies with human CBP and rat p22 proteins. CBP regulates a protein (NHE 1) which activates cell proliferation, differentiation and neoplastic transformation. p22 is thought to be a regulator of constitutive endocytotic events. CAB is expressed in proliferating tissue and also in reproductive, haemopoietic/immune, cardiovascular, gastrointestinal, brain, bladder and kidney tissue. CAB therefore appears to play a role in apoptotic, neoplastic, immunological and reproductive disorders. Apoptotic disorders include acquired immunodeficiency syndrome (AIDS), Alzheimer's disease and Parkinson's disease. Neoplastic disorders include leukaemia, lymphoma, melanoma and myeloma. Immunological disorders include AIDS, Addison's disease and anaemia. Reproductive disorders include infertility, prolactin production and disruptions of the oestrous cycle. Antibodies that specifically bind to CAB are useful for diagnosis of disorders associated with CAB, or in assays to monitor treatment involving (ant)agonists and/or CAB inhibitors.

Sequence 214 AA;

Query Match 66.7%; Score 42; DB 20; Length 214;
 Best Local Similarity 80.0%; Pred. No. 3;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 4 DSDGTITIEE 13

Db 125 dsgdgritlee 134

RESULT 3

R14150

ID R14150 standard; Protein; 461 AA.

XX R14150;

XX R14150;

XX 09-DEC-1991 (first entry)

XX MSP encoded by pUCRS (DSM 5803).

XX Major secretion product; expression cassette.

XX

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: November 8, 2000, 08:51:37 ; Search time 138.73 Seconds
(without alignments)
3.204 Million cell updates/sec

Title: US-09-236-468a-2_COPY_24_36
Perfect score: 63
Sequence: 1 AQLSDSGTITIEE 13
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 268485 seqs, 34193795 residues
Total number of hits satisfying chosen parameters: 268485
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36.*
1: /SIDSL1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	ID	Description	
1	63	100.0	541	18	W12695	G-protein parathyr
2	42	66.7	214	20	Y31625	Human calcineurin
3	42	66.7	461	12	R14150	MSP encoded by pUC
4	42	66.7	461	12	R14530	Usp45 protein. La
5	41	65.1	344	21	Y52206	Mycobacterium tube
6	39	61.9	243	18	W21782	Protein encoded by
7	39	61.9	4928	20	Y39300	SpnD a polyketide
8	37	58.7	149	11	R04583	Recombinant calmod
9	37	58.7	149	16	R78523	Human calmodulin p
10	37	58.7	149	21	Y50797	Human calmodulin/N
11	37	58.7	151	9	P80162	Biosynthetic multi
12	37	58.7	375	16	R78526	GST-calmodulin fus

13	37	58.7	409	16	R78525	Protein A-calmodul
14	37	58.7	416	20	W92993	S. aureus fsty pro
15	37	58.7	535	16	R78524	Maltose binding pr
16	37	58.7	544	20	Y23905	Amino acid sequenc
17	37	58.7	545	20	Y23904	Streptococcus pyog
18	37	58.7	642	19	W71645	Fluorescent calmod
19	37	58.7	642	19	W71647	Fluorescent calmod
20	37	58.7	652	19	W71646	Fluorescent calmod
21	37	58.7	656	19	W71648	Human interferon r
22	36	57.1	191	19	W64199	Human protein phos
23	36	57.1	191	19	W62287	Amino acid sequenc
24	36	57.1	191	19	W51215	Feline immunoglob
25	36	57.1	333	12	R11014	Sequence region of
26	36	57.1	333	14	R48367	Constant encoded b
27	36	57.1	333	15	R54094	Fungal signal reco
28	36	57.1	534	17	R91309	Arabidopsis thalia
29	36	57.1	539	21	Y99471	Heat shock protein
30	36	57.1	541	19	W77406	Streptococcus pneu
31	36	57.1	541	20	Y23902	Amino acid sequenc
32	36	57.1	551	20	Y23907	Human SH3DIA prote
33	36	57.1	648	20	Y32157	Human SH3DIA prote
34	36	57.1	1144	20	Y32154	Human SH3DIA prote
35	36	57.1	1214	21	Y57444	Mouse Esel protein
36	36	57.1	1215	20	Y32156	Human SH3DIA prote
37	36	57.1	1220	20	Y32155	Human SH3DIA prote
38	36	57.1	1715	21	Y57449	Mouse Esell protei
39	35	55.6	22	16	R85567	Calmodulin C-termi
40	35	55.6	136	21	Y77954	A. thaliana enviro
41	35	55.6	146	21	Y87548	Human disease-asso
42	35	55.6	149	17	R81571	Potato calmodulin.
43	35	55.6	206	20	Y25651	Betula sp. allerge
44	35	55.6	276	20	W89871	Antigen from clust
45	35	55.6	276	20	W89994	Expressed antigen

ALIGNMENTS

RESULT 1	
W12695	
ID	W12695 standard; Protein; 541 AA.
XX	
AC	W12695;
XX	
DT	31-MAY-1997 (first entry)
XX	
DE	G-protein parathyroid hormone receptor HLTDG74.
XX	
KW	G-protein parathyroid hormone receptor; HLTDG74; parathormone; PTH;
KW	calcium; signal transduction; agonist; antagonist; hypocalcaemia;
KW	hyperphosphataemia; hypoparathyroidism; chronic tetany;
KW	osteoporosis; hypercalcaemia; hypoparathyroidism; hypophosphataemia;
KW	kidney stone; nephrolisis; therapy; diagnosis.
OS	Homo sapiens.
XX	
PN	WO9639433-A1.
XX	
PD	12-DEC-1996.
XX	
PF	05-JUN-1995; 95WO-US07085.
XX	
PR	05-JUN-1995; 95WO-US07085.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Li Y, Rosen CA, Ruben SM, Soppet DR;
XX	
DR	WPI: 1997-043068/04.
XX	
NR	N-PSDB; T59619.
XX	
PT	Human G-protein parathyroid hormone receptor, HLTDG74 - used to identify (ant)agonists, used in the treatment of hypo- or

QY 1 IMQDDPQNSIATSVKSK 19
: |||: : |||:
Db 134 VNDPDKKAILKTNADKTK 152

RESULT 13
O27875
ID Q27875 PRELIMINARY; PRT; 200 AA.
AC Q27875;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)
DE DEOXYCYTIDINE TRIPHOSPHATE DEAMINASE.
GN MTH1847.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RX MEDLINE; 98037514.
RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., DuBois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Nollung J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT delta: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
DR EMBL; AF000937; AAB86313.1; -;
DR INTERPRO; IPR001428; -;
DR PFAM; PF00692; dUMPase; 1.
SQ SEQUENCE 200 AA; 22818 MW; E228E4CA9CB9B7E1 CRC64;

Query Match 40.6%; Score 41; DB 1; Length 200;
Best Local Similarity 53.8%; Pred. No. 29;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 DDPQNSIATSVKSK 16
: |||: : |||:
Db 27 DDPERQIPSSVD 39

RESULT 14
O13955
ID O13955 PRELIMINARY; PRT; 905 AA.
AC O13955;
DT 01-JUN-1998 (TEMBLrel. 06, Created)
DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)
DT 01-JUN-2000 (TEMBLrel. 14, Last annotation update)
DE HYPOTHETICAL 103.7 KDA PROTEIN C23H4.14 IN CHROMOSOME I.
GN SPAC23H4.14.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
OC Schizosaccharomycetaceae; Schizosaccharomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Brown D., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z98977; CAB11668.1; -;
DR INTERPRO; IPR000345; -;
DR INTERPRO; IPR001180; -;
DR INTERPRO; IPR001220; -;
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN_1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
KW Hypothetical protein.
FT DOMAIN 46 49 POLY-GLU.
SQ SEQUENCE 905 AA; 103695 MW; 8C0FD258D4745D0E CRC64;

Query Match 40.6%; Score 41; DB 3; Length 905;
Best Local Similarity 37.5%; Pred. No. 1.4e+02;
Matches 9; Conservative 5; Mismatches 2; Indels 8; Gaps 1;

QY 2 MQDDPQNSI-----EATSVK 17
: |||: : |||:
Db 602 LSEDPQNSIVIFIDENSEASTISK 625

RESULT 15
P78953
ID P78953 PRELIMINARY; PRT; 920 AA.
AC P78953;
DT 01-MAY-1997 (TEMBLrel. 03, Created)
DT 01-MAY-1997 (TEMBLrel. 03, Last sequence update)
DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)
DE DMF1 GENE.
GN DMF1.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
OC Schizosaccharomycetaceae; Schizosaccharomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RX MEDLINE; 97102672.
RA Sohrmann M., Fankhauser C., Brodbeck C., Simanis V.;
RT "The dmfl/midl gene is essential for correct positioning of the
RT division septum in fission yeast.";
RL Genes Dev. 10:2707-2720(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Seeger K., Harris D., McDougall R.C., Rajandream M.A., Barrell B.G.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y07599; CAA68873.1; -;
DR EMBL; AL132870; CAB60689.1; -;
DR INTERPRO; IPR001849; -;
DR PFAM; PF00169; PH; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
SQ SEQUENCE 920 AA; 102366 MW; CA95FB48A20549B5 CRC64;

Query Match 40.6%; Score 41; DB 3; Length 920;
Best Local Similarity 42.1%; Pred. No. 1.5e+02;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 IMQDDPQNSIATSVKSK 19
: |||: : |||:
Db 378 LQENMDEISATSIKSK 396

Search completed: November 8, 2000, 08:56:07
Job time: 518 sec

SO SEQUENCE 538 AA; 59445 MW; C342198841B87CCA CRC64;

Query Match 41.1%; Score 41.5; DB 10; Length 538;
Best Local Similarity 36.0%; Pred. No. 68;
Matches 9; Conservative 5; Mismatches 4; Indels 7; Gaps 1;

Qy 1 IMODDPQNSIEAT-----SVDRS 18

Db : ||||| : ||||| : ||||| : |||||

88 VYSDDPQAQLEATTQFRKLIIERS 112

RESULT 10

Q9UU01

ID Q9UU01 PRELIMINARY; PRT; 143 AA.

AC Q9UU01;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)

DE SEXUAL DIFFERENTIATION PROCESS PROTEIN (FRAGMENT).

GN ISP4+.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;

CC Schizosaccharomycetaceae; Schizosaccharomycetes.

FN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-968 H90;

RA Ding D.;

RT "Generation and analysis of GFP-gene fusion library of fission

yeast.";

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB027895; BAA87199.1; -.

FT NON_TER 143

FT SEQUENCE 143 AA; 15898 MW; 285FECE92081FE11 CRC64;

Query Match 40.6%; Score 41; DB 3; Length 143;

Best Local Similarity 47.4%; Pred. No. 20;

Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 2 MODDPQNSIEATSVDRKSOY 20

Db : ||||| : ||||| : ||||| : |||||

14 MNDSPSTKEKADSVDISDY 32

RESULT 11

Q9VZ43

ID Q9VZ43 PRELIMINARY; PRT; 187 AA.

AC Q9VZ43;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)

DE CGI5204 PROTEIN.

GN CGI5204.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

CC Ephydroidea; Drosophilidae; Drosophila.

FN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RA MEDLINE; 20196006.

RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,

George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003484; AAF47985.1; -.
DR FLYBASE; FBgn0030256; CGI5204.
SQ SEQUENCE 187 AA; 20511 MW; C7A4502FF0B0C533 CRC64;

Query Match 40.6%; Score 41; DB 5; Length 187;

Best Local Similarity 38.9%; Pred. No. 27;

Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 3 QDDPONSIEATSVDRKSOY 20

Db : ||||| : ||||| : ||||| : |||||

89 QDDLQALDSDSNSTDY 106

RESULT 12

Q45030

ID Q45030 PRELIMINARY; PRT; 194 AA.

AC Q45030;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)

DE PLE OSPC GENE (FRAGMENT).

GN OSPA.

OS Borrelia burgdorferi (Lyme disease spirochete).

OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

FN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PLE;

RA MEDLINE; 95395018.

RA Jauris-Heipke S., Liegl G., Preac-Mursic V., Roessler D., Schwab E.,

RA Soutschek E., Will G., Wilske B.;

RT "Molecular analysis of genes encoding outer surface protein C (OspC)

of Borrelia burgdorferi sensu lato: relationship to ospA genotype and

evidence of lateral gene exchange of ospC.";

J. Clin. Microbiol. 33:1860-1866(1995).

RL EMBL; X80255; CAA56548.1; -.

DR INTERPRO; IPR001800; -.

DR PFAM; PF01441; Lipoprotein_6; 1.

FT NON_TER 194

FT NON_TER 194

SQ SEQUENCE 194 AA; 20523 MW; 11D409DBBFD23288 CRC64;

Query Match 40.6%; Score 41; DB 2; Length 194;

Best Local Similarity 36.8%; Pred. No. 28;

Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY.
 RX MEDLINE; 20196006.
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berland B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler J., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003686; AAF54522.1; -.
 DR FLYBASE; FBgn0037810; CG12819.
 SQ SEQUENCE 920 AA; 103007 MW; 0F56010F3DC8CAE0 CRC64;

Query Match 41.6%; Score 42; DB 5; Length 920;
 Best Local Similarity 57.1%; Pred. No. 99;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 MQDDPQNSIATSV 15
 Db 473 MDDKQNKIDVTEV 486
 | | | | | | | | | |

RESULT 8
 O49602 PRELIMINARY; PRT; 528 AA.
 AC O49602;

DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE IMPORTIN ALPHA-LIKE PROTEIN (FRAGMENT).
 GN IMPA-4.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; Rosidae; eurosids II; Brassicales;
 OC Brassicaceae; Arabidopsis.

RN SEQUENCE FROM N.A.
 RC TISSUE=VEGETATIVE TISSUE;
 RA Schledz M., Leclerc D., Neuhaus G., Merkle T.;
 RL Plant Physiol. 116:868-868(1998).
 DR EMBL; Y14616; CAA74966.1; -.
 DR HSSP; Q02821; IBK5.
 DR MENDEL; 24707; Arath; 2729; 24707.
 DR INTERPRO; IPR000225; -.
 DR INTERPRO; IPR000495; -.
 DR INTERPRO; IPR002652; -.
 DR PRAM; PF00514; Armadillo_seg; 8.
 DR PRAM; PF01749; IBB; 1.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 528 AA; 58265 MW; D794CALCE959E731 CRC64;
 Query Match 41.1%; Score 41.5; DB 10; Length 528;
 Best Local Similarity 36.0%; Pred. No. 66;
 Matches 9; Conservative 5; Mismatches 4; Indels 7; Gaps 1;

QY 1 IMQDDPQNSIEAT-----SVDKS 18
 : | | | | : | | | | : | | | |
 Db 79 VYSDPQAQLEATTQFRKLLSIERS 103

RESULT 9
 O80480 PRELIMINARY; PRT; 538 AA.
 ID O80480
 AC O80480
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE T12M4.2 PROTEIN.
 GN T12M4.2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; Rosidae; eurosids II; Brassicales;
 OC Brassicaceae; Arabidopsis.

[1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Oji, O, Kwan, A,
 RA Liu S., Li J., Araujo R., Au M., Brendel V., Buehler E., Conway A.B.,
 RA Conway A.R., Dewar K., Feng J., Kim C., Kurtz D., Li Y., Palm C.J.,
 RA Shinn P., Sun H., Davis R.W., Ecker J.R., Federspiel N.A.,
 RA Theologis A.;
 RA "Arabidopsis thaliana chromosome 1 BAC T12M4 sequence, complete
 sequence."
 RT Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Theologis A.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Theologis A.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 [4]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Theologis A.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC003114; AAC24079.1; -.
 DR HSSP; Q02821; IBK5.
 DR MENDEL; 31141; Arath; 2729; 31141.
 DR INTERPRO; IPR000225; -.
 DR INTERPRO; IPR002652; -.
 DR PRAM; PF00514; Armadillo_seg; 8.
 DR PRAM; PF01749; IBB; 1.

RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA	Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA	Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA	Wang Z.-Y., Wasarman D.A., Weinstein G.M., Weissenbach J.,
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA	Ye J., Yeh R.-F., Zavari J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT	"The genome sequence of <i>Drosophila melanogaster</i> .";
RL	Science 287:2185-2195(2000).
RN	EMBL; AE003840; AAF59211.1; -.
DR	FLYBASE; FBgn0033206; CG12042.
SO	SEQUENCE 514 AA; 57925 MW; 61424012AA831E95 CRC64;
Query Match	41.6%; Score 42; DB 5; Length 514;
Best Local Similarity	56.2%; Pred. No. 53;
Matches	9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Qy	4 DDPNSTSEATSVKQS 19 :: :
Db	361 EDPOLSTEATRVAKEE 376 :: :
RESULT 6	
O42943	PRELIMINARY; PRT; 618 AA.
ID O42943	AC 012943;
DT 01-JUN-1998	(TrEMBLrel. 06, Created)
DD 01-JUN-1998	(TrEMBLrel. 06, Last sequence update)
DE 01-JUN-2000	(TrEMBLrel. 14, Last annotation update)
DE PROBABLE ATP-DEPENDENT TRANSPORTER C16H5.08C.	
GN SPEC16H5.08C.	
OS Schizosaccharomyces pombe (Fission yeast).	
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;	
OC Schizosaccharomycetaceae; Schizosaccharomycetes.	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=972;	
RA Wood V., Rajandream M.A., Barrell B.G., Skelton J., Churcher C.M.;	
RL Submitted (MAR-1998), to the EMBL/GenBank/DDBJ databases.	
CC - SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY	
CC (ABC TRANSPORTERS). EF-3 SUBFAMILY.	
DR EMBL; AL022104; CAAI7906.1; -.	
DR INTERPRO; IPR001617; -.	
DR PFAM; PF00005; ABC_tran; 2.	
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.	
KW Hypothetical protein; ATP-binding; Transport; Repeat.	
FT NP_BIND 108 115 ATP (POTENTIAL).	
FT NP_BIND 423 430 ATP (POTENTIAL).	
SO SEQUENCE 618 AA; 69212 MW; 2f547d71a7986CA4 CRC64;	
Query Match	41.6%; Score 42; DB 3; Length 618;
Best Local Similarity	37.3%; Pred. No. 65;
Matches	6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
Qy	1 IMQDPONSIEATSVD 16 :: : : : :
Db	534 LLLEPTNHLDTSID 549 :: : : : :
RESULT 7	
Q9VGZ9	PRELIMINARY; PRT; 920 AA.
ID Q9VGZ9	AC 09VGZ9;
DT 01-MAY-2000	(TrEMBLrel. 13, Created)
DD 01-MAY-2000	(TrEMBLrel. 13, Last sequence update)
DE 01-MAY-2000	(TrEMBLrel. 13, Last annotation update)
DE CG12819 PROTEIN.	
GN CG12819.	


```

Best Local Similarity 56.2%; Pred. No. 45;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 MQDDPQNSIEATSVDK 17
: ||| ||| | |||
Db 1187 LPDDPHNSIGAMVDQ 1202

RESULT 2
ID O01938 PRELIMINARY; PRT; 1876 AA.
AC O01938; Q24453;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE PHOSPHOINOSITIDE 3-KINASE.
GN PI3K 68D.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 96362138.
RA MacDougall L.K., Domin J., Waterfield M.D.;
RT "A family of phosphoinositide 3-kinases in Drosophila identifies a new
RL Curr. Biol. 5:1404-1415(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA MacDougall L.K.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: X92892; CAA63485.1; -.
DR HSSP: P21707; IRSY.
DR FLYBASE: FBgn0015278; PI3K68D.
DR INTERPRO: IPR000008; -.
DR INTERPRO: IPR000341; -.
DR INTERPRO: IPR000403; -.
DR INTERPRO: IPR001263; -.
DR INTERPRO: IPR001683; -.
DR INTERPRO: IPR002420; -.
DR PFAM: PF00168; C2; 1.
DR PFAM: PF00454; PI3_PI4_kinase; 1.
DR PFAM: PF00613; PI3Ka; 1.
DR PFAM: PF00787; PX; 1.
DR PFAM: PF00792; PI3K_C2; 1.
DR PFAM: PF00794; PI3K_rbd; 1.
DR PROSITE: PS00916; PI3_4_KINASE_2; 1.
DR PROSITE: PS50004; C2_DOMAIN_2; 1.
DR PROSITE: PS50004; C2_DOMAIN_2; 2.
SQ SEQUENCE 1876 AA; 210488 MW; 72D07CB9F9E5B901 CRC64;

Query Match 45.5%; Score 46; DB 5; Length 1876;
Best Local Similarity 56.2%; Pred. No. 45;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 MQDDPQNSIEATSVDK 17
: ||| ||| | |||
Db 1187 LPDDPHNSIGAMVDQ 1202

RESULT 3
Q9VTN5 PRELIMINARY; PRT; 1876 AA.
AC Q9VTN5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE PI3K68D PROTEIN.
GN PI3K68D.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

```

```

OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 20196006.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abiril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003543; AAF50011.1; -.
DR FLYBASE: FBgn0015278; PI3K68D.
DR INTERPRO: IPR000008; -.
DR INTERPRO: IPR000341; -.
DR INTERPRO: IPR000403; -.
DR INTERPRO: IPR001263; -.
DR INTERPRO: IPR001683; -.
DR INTERPRO: IPR002420; -.
DR PFAM: PF00168; C2; 1.
DR PFAM: PF00454; PI3_PI4_kinase; 1.
DR PFAM: PF00613; PI3Ka; 1.
DR PFAM: PF00787; PX; 1.
DR PFAM: PF00792; PI3K_C2; 1.
DR PFAM: PF00794; PI3K_rbd; 1.
DR PROSITE: PS00916; PI3_4_KINASE_2; 1.
DR PROSITE: PS50004; C2_DOMAIN_2; 2.
SQ SEQUENCE 1876 AA; 210473 MW; BD1931D659A6B983 CRC64;

Query Match 45.5%; Score 46; DB 5; Length 1876;
Best Local Similarity 56.2%; Pred. No. 45;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 MQDDPQNSIEATSVDK 17
: ||| ||| | |||
Db 1187 LPDDPHNSIGAMVDQ 1202

RESULT 4
Q9V0T8

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:56:04 : Search time 152.43 Seconds
(without alignments)
12.251 Million cell updates/sec

Title: US-09-236-468a-2_COPY_214_233

Perfect score: 101

Sequence: 1 IMQDDPQNSIATSDKSQY 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 297973

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_14:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	46	45.5	1876	5 Q24209	Q24209 drosophila
2	46	45.5	1876	5 Q01938	Q01938 drosophila
3	46	45.5	1876	5 Q9VTN5	Q9VTN5 drosophila
4	43	42.6	394	1 Q9V0T8	Q9V0T8 pyrococcus
5	42	41.6	514	5 Q9V4Q4	Q9V4Q4 drosophila
6	42	41.6	618	3 Q42943	Q42943 schizosacch
7	42	41.6	920	5 Q9VG29	Q9VG29 drosophila
8	41.5	41.1	528	10 Q49602	Q49602 arabidopsis
9	41.5	41.1	538	10 Q80480	Q80480 arabidopsis
10	41	40.6	143	3 Q9U001	Q9U001 schizosacch
11	41	40.6	187	5 Q9V243	Q9V243 drosophila
12	41	40.6	194	2 Q45030	Q45030 borrelia bu
13	41	40.6	200	1 Q27875	Q27875 methanobact
14	41	40.6	905	3 Q13955	Q13955 schizosacch
15	41	40.6	920	3 P78953	P78953 schizosacch
16	40.5	40.1	887	11 Q62940	Q62940 rattus norv
17	40	39.6	66	2 Q92763	Q92763 chlamydia p
18	40	39.6	241	10 Q39778	Q39778 gossypium h
19	40	39.6	246	10 Q42488	Q42488 gossypium b

20	40	39.6	268	11 Q9QVN6	Q9QVN6 mus sp. b-c
21	40	39.6	316	2 Q9RK71	Q9RK71 streptomyce
22	40	39.6	325	5 Q02322	Q02322 caenorhabdi
23	40	39.6	329	2 Q86190	Q86190 erwinia chr
24	40	39.6	530	10 P93759	P93759 arabidopsis
25	40	39.6	609	5 Q19719	Q19719 caenorhabdi
26	40	39.6	823	4 Q9UPT2	Q9UPT2 homo sapien
27	40	39.6	843	11 Q62765	Q62765 rattus norv
28	40	39.6	931	10 Q04026	Q04026 arabidopsis
29	40	39.6	1130	5 Q9XTI9	Q9XTI9 caenorhabdi
30	40	39.6	1164	10 Q9XIK6	Q9XIK6 arabidopsis
31	39.5	39.1	136	8 P92865	P92865 sathon falc
32	39.5	39.1	653	5 Q97245	Q97245 plasmodium
33	39	38.6	170	5 Q94563	Q94563 entamoeba h
34	39	38.6	194	5 Q07592	Q07592 entamoeba h
35	39	38.6	233	3 Q9USA9	Q9USA9 schizosacch
36	39	38.6	238	5 Q07593	Q07593 entamoeba h
37	39	38.6	242	5 Q07594	Q07594 entamoeba h
38	39	38.6	251	2 Q9XAQ6	Q9XAQ6 streptomyce
39	39	38.6	291	2 Q55426	Q55426 synechocyst
40	39	38.6	496	1 Q27785	Q27785 methanobact
41	39	38.6	514	2 Q45087	Q45087 burkholderi
42	39	38.6	518	5 Q9VF72	Q9VF72 drosophila
43	39	38.6	580	11 Q9QXY4	Q9QXY4 rattus norv
44	39	38.6	606	5 Q19731	Q19731 caenorhabdi
45	39	38.6	617	5 Q9U3G1	Q9U3G1 caenorhabdi

ALIGNMENTS

RESULT 1

Q24209 ID Q24209 PRELIMINARY; PRT; 1876 AA.

AC Q24209;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DE 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)

DE PHOSPHOINOSITIDE 3-KINASE.

GN CPK.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=OREGON R;

RX MEDLINE; 96278830.

RA Molz L.M., Chen Y.W., Hirano M., Williams L.T.;

RT "Cpk is a novel class of Drosophila PtdIns 3-kinase containing a C2 domain."

RL J. Biol. Chem. 271:13892-13899(1996).

DR EMBL; U52192; AAC47117.1; -.

DR HSSP; P21707; IRSY.

DR FLYBASE; FBgn0015278; P13K68D.

DR INTERPRO; IPR000008; -.

DR INTERPRO; IPR000341; -.

DR INTERPRO; IPR000403; -.

DR INTERPRO; IPR001263; -.

DR INTERPRO; IPR001683; -.

DR INTERPRO; IPR002420; -.

DR PFAM; PF00168; C2; 1.

DR PFAM; PF00454; P13_PI4_kinase; 1.

DR PFAM; PF00613; P13Ka; 1.

DR PFAM; PF00787; PX; 1.

DR PFAM; PF00792; P13K_C2; 1.

DR PFAM; PF00794; P13K_rbd; 1.

DR PROSITE; PS00916; P13_4_KINASE_2; 1.

DR PROSITE; PS00004; C2_DOMAIN_2; 1.

SQ SEQUENCE 1876 AA; 210504 MW; BCA64ADA464C2763 CRC64;

Query Match

45.5%; Score 46; DB 5; Length 1876;

THIS PAGE BLANK (USPTO)

```
RESULT 14
143B_TOBAC
ID 143B_TOBAC STANDARD; PRT; 255 AA.
AC O49995;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE 14-3-3-LIKE PROTEIN B.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Nicotiana.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98105599.
RA Piotrowski M., Oecking C.;
RT "Five new 14-3-3 isoforms from Nicotiana tabacum L.: implications for
the phylogeny of plant 14-3-3 proteins.";
RL planta 204:127-130(1998).
CC -1- SIMILARITY: BELONGS TO THE 14-3-3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U91723; AAC49891.1; -.
DR INTERPRO: IPR000308; -.
DR PFAM; PF00244; 14-3-3; 1.
DR PRINTS; PR00305; 1433ZETA.
DR PROSITE; PS00796; 1433_1; 1.
DR PROSITE; PS00797; 1433_2; 1.
KW Multigene family.
SQ SEQUENCE 255 AA; 28827 MW; 65CE3CD2FD83DF1E CRC64;

Query Match 39.6%; Score 40; DB 1; Length 255;
Best Local Similarity 44.4%; Pred. No. 17;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 2 MQDDPQNSIEATSDVKSQ 19
| | | : | : | |
Db 237 MQDDGADEIKETKTDEQ 254

RESULT 15
OBFL_MOUSE
ID OBFL_MOUSE STANDARD; PRT; 256 AA.
AC Q64693;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE POU DOMAIN CLASS 2, ASSOCIATING FACTOR 1 (B-CELL-SPECIFIC
DE COACTIVATOR OBF-1) (OCT BINDING FACTOR 1) (BOB1) (OCA-B).
GN POU2AF1 OR OBF-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE; 96102041.
RA Pfisterer P., Zwilling S., Hess J., Wirth T.;
RT "Functional characterization of the murine homolog of the B cell-
RL J. Biol. Chem. 270:29870-29880(1995).
RN [2]
RP SEQUENCE FROM N.A.
```

CC UPSTREAM ACTIVATION SEQUENCE OF GLN1, THE GENE ENCODING GLUTAMINE
 CC SYNTHETASE. URE2 MAY CATALYTICALLY INACTIVATE GLN3 IN RESPONSE TO
 CC AN INCREASE IN THE INTRACELLULAR CONCENTRATION OF GLUTAMINE.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- SIMILARITY: BELONGS TO THE GATA-TYPE TRANSCRIPTION FACTOR FAMILY.
 CC HIGH, TO OTHER FUNGAL NITROGEN REGULATORY PROTEINS.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; M35267; AAA34645.1; -
 CC EMBL; U18796; AAB64575.1; -
 CC DR PIR; S22280; S22280.
 CC DR HSP; P17429; SGAT.
 CC SGD; S0000842; GLN3.
 CC DR INTERPRO; IPR000679; -
 CC DR PFAM; PF00320; GATA; 1.
 CC DR PRINTS; PR00619; GATAZFINGER.
 CC DR PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
 CC DR PROSITE; PS00114; GATA_ZN_FINGER_2; 1.
 CC KW Transcription regulation; Activator; DNA-binding; Zinc-finger;
 CC Nuclear protein; Nitrate assimilation.
 CC FT ZN_FING 306 330 GATA-TYPE
 CC FT DOMAIN 351 361 ARG/LYS-RICH (BASIC).
 CC FT CONFLICT 474 474 P -> G (IN REF. 1).
 CC SQ SEQUENCE 730 AA; 79382 MW; 3159E184469942E CRC64;

Query Match 40.6%; Score 41; DB 1; Length 730;
 Best Local Similarity 87.5%; Pred. No. 41;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MQDDPONS 9
 Db 1 MQDDPENS 8

RESULT 12

ISP4_SCHPO
 ID ISP4_SCHPO STANDARD; PRT; 776 AA.
 AC P40900;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE SEXUAL DIFFERENTIATION PROCESS PROTEIN ISP4.
 GN ISP4.
 OS Schizosaccharomyces pombe (Fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
 CC Schizosaccharomycetaceae; Schizosaccharomyces.
 [1]
 RP SEQUENCE FROM N.A.
 RN MEDLINE; 95042833.
 RA Sato S., Suzuki H., Widyastuti U., Hotta Y., Tabata S.;
 RT "Identification and characterization of genes induced during sexual
 RT differentiation in Schizosaccharomyces pombe."
 RL Curr. Genet. 26:31-37(1994).
 CC -1- DEVELOPMENTAL STAGE: TRANSCRIBED SPECIFICALLY DURING SEXUAL
 CC DEVELOPMENT

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; D14061; BAA03147.1; -

DR PIR; S43741; S43741.
 DR PIR; S45495; S45495.
 SQ SEQUENCE 776 AA; 88256 MW; 0BEE00B9C6C3725 CRC64;

Query Match 40.6%; Score 41; DB 1; Length 776;
 Best Local Similarity 47.4%; Pred. No. 44;
 Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 MQDDPONSIEATSVDKSOY 20
 Db 14 MNDSPSTKERADSVDISDY 32

RESULT 13

E6_GOSHI
 ID E6_GOSHI STANDARD; PRT; 238 AA.
 AC Q01197;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE PROTEIN E6.
 GN E6.
 OS Gossypium hirsutum (Upland cotton).
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 CC Malvales; Malvaceae; Gossypium.
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CV. COKER 312; TISSUE=FIBER;
 RX MEDLINE; 92335179.
 RA John M.E., Crow L.J.;
 RT "Gene expression in cotton (Gossypium hirsutum L.) fiber: cloning of
 RT the mRNAs."
 RL Proc. Natl. Acad. Sci. U.S.A. 89:5769-5773(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COKER 312;
 RX MEDLINE; 96178868.
 RA John M.E.;
 RT "Structural characterization of genes corresponding to cotton fiber
 RT mRNA, E6; reduced E6 protein in transgenic plants by antisense
 RT gene."
 RL Plant Mol. Biol. 30:297-306(1996).
 CC -1- TISSUE SPECIFICITY: IT IS PREDOMINANTLY EXPRESSED IN FIBER CELLS.
 CC -1- DEVELOPMENTAL STAGE: CONCENTRATION OF E6 IS HIGHEST DURING THE
 CC LATE PRIMARY CELL WALL AND EARLY CELL WALL SYNTHESIS STAGES.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1, MET-5, MET-10 OR MET-21 IS
 CC THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; M92051; AAA33055.1; -
 CC EMBL; M92051; AAA33056.1; ALT_INIT.
 CC DR EMBL; U30505; AAB03079.1; -
 CC DR PIR; A46130; A46130.
 KW Cell wall.
 SQ SEQUENCE 238 AA; 28226 MW; 7162101A9A91A4FB CRC64;

Query Match 39.6%; Score 40; DB 1; Length 238;
 Best Local Similarity 47.1%; Pred. No. 16;
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 4 DDQNSIEATSVDKSOY 20
 Db 107 DEPNSIPSSNNKDTY 123

RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., DuBois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT deltaH: functional analysis and comparative genomics.";
 RL J. Bacteriol. 179:7135-7155(1997).
 CC -1- CATALYTIC ACTIVITY: DCTP + H(2)O -> DUTP + NH(3).
 CC -1- SIMILARITY: BELONGS TO THE DCTP DEAMINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000937; AAB86313.1; ALT_INIT.
 DR PFAM: PF00602; dUTPase; 1.
 KW Hydrolase.
 SQ SEQUENCE 197 AA; 22426 MW; 96A019F9DBA3528F CRC64;

Query Match 40.68; Score 41; DB 1; Length 197;
 Best Local Similarity 53.88; Pred. No. 8.8;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 DDQNSIEATSVSD 16
 ||: ||: |||
 Db 24 DDPERQTPSVSD 36

RESULT 10
 PEDD_PEDAC STANDARD; PRT; 724 AA.
 AC P36497;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE PEDIOCIN PA-1 TRANSPORT ATP-BINDING PROTEIN PEDD (PEDIOCIN ACH
 DE TRANSPORT ATP-BINDING PROTEIN PAPD).
 GN PEDD OR PAPD.
 OS Pediococcus acidilactici.
 OC Plasmid pSRQ11, and Plasmid pSMB74.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
 OC Pediococcus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PAC-1.0; PLASMID-PSRQ11;
 RX MEDLINE; 92384551.
 RA Marugg J.D., Gonzalez C.F., Kunka B.S., Ledebor A.M., Pucci M.J.,
 RA Toonen M.Y., Walker S.A., Zoetmulder L.C.M., Vandenbergh P.A.;
 RT "Cloning, expression, and nucleotide sequence of genes involved in
 RT production of pediocin PA-1, and bacteriocin from *Pediococcus*
 RT acidilactici PAC1.0";
 RL Appl. Environ. Microbiol. 58:2360-2367(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H; PLASMID-PSMB74;
 RX MEDLINE; 94288967.
 RA Motilagh A.M., Bukhtiyarova M.B., Ray B.R.;
 RT "Complete nucleotide sequence of pSMB 74, a plasmid encoding the
 RT production of pediocin ACH in *Pediococcus acidilactici*.";
 RL Lett. Appl. Microbiol. 18:305-312(1994).
 CC -1- FUNCTION: INVOLVED IN THE EXPORT PROCESS OF THE BACTERIOCIN
 CC PEDIOCIN PA-1/ACH. IS ALSO ESSENTIAL FOR PEDIOCIN PRODUCTION.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS). HLYB SUBFAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M83924; AAA25561.1; -;
 DR EMBL: U02482; AAC43296.1; -;
 DR PIR: D48941; D48941.
 DR HSP: P13569; INED.
 DR INTERPRO: IPR001140; -;
 DR INTERPRO: IPR001617; -;
 DR PFAM: PF00664; ABC_membrane; 1.
 DR PFAM: PF00005; ABC_tran; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KW Plasmid; Transmembrane; Transport; ATP-binding.
 FT TRANSMEM 170 186 POTENTIAL.
 FT TRANSMEM 207 222 POTENTIAL.
 FT TRANSMEM 232 248 POTENTIAL.
 FT TRANSMEM 286 300 POTENTIAL.
 FT TRANSMEM 309 325 POTENTIAL.
 FT TRANSMEM 402 417 POTENTIAL.
 FT NP_BIND 519 526 ATP (POTENTIAL).
 SQ SEQUENCE 724 AA; 81651 MW; B897C680EB519AF5 CRC64;

Query Match 40.68; Score 41; DB 1; Length 724;
 Best Local Similarity 37.58; Pred. No. 40;
 Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 5 DPQNSIEATSVDSQY 20
 ||: ||: ||: |||
 Db 114 DPDPVTVRTTKISQSF 129

RESULT 11
 GLN3_YEAST STANDARD; PRT; 730 AA.
 AC P18494;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE NITROGEN REGULATORY PROTEIN GLN3.
 GN GLN3 OR YER040W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomycetes.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 92049353.
 RA Minehart P.L., Magasanik B.;
 RT "Sequence and expression of GLN3, a positive nitrogen regulatory gene
 RT of *Saccharomyces cerevisiae* encoding a protein with a putative zinc
 RT finger DNA-binding domain.";
 RL Mol. Cell. Biol. 11:6216-6228(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
 RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
 RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
 RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
 RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
 RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: POSITIVE NITROGEN REGULATORY PROTEIN. REQUIRED FOR
 CC THE ACTIVATION OF TRANSCRIPTION OF A NUMBER OF GENES (INCLUDING
 CC THE ALLANTOIN PATHWAY GENES) IN RESPONSE TO THE REPLACEMENT OF
 CC GLUTAMINE BY GLUTAMATE AS SOURCE OF NITROGEN. BINDS THE NITROGEN

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; L35572; AAA62162.1; -;
DR EMBL; U61155; AAC52842.1; -;
DR HSP; P06601; IFJL.
DR MGD; MGI:107792; LHX5.
DR INTERPRO; IPR001356; -;
DR INTERPRO; IPR001781; -;
DR PFAM; PF00412; LIM; 2.
DR PFAM; PF00046; homeobox; 1.
DR PROSITE; PS00478; LIM_DOMAIN_1; 2.
DR PROSITE; PS00023; LIM_DOMAIN_2; 2.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Nuclear protein; Repeat; LIM motif;
KW Metal-binding; Zinc; Transcription regulation.
FT DOMAIN 5 55 LIM.
FT DOMAIN 64 118 LIM.
FT DNA_BIND 180 239 HOMEBOX.
SQ SEQUENCE 402 AA; 44387 MW; 8F2DBBF1C6237227 CRC64;

Query Match 43.6%; Score 44; DB 1; Length 402;
Best Local Similarity 50.0%; Pred. No. 6.4;
Matches 8; Conservative 4; Mismatches 4; Indels 4; Gaps 0;

QY 2 MQDDPQNSIETSVDK 17
:||||: :|||
Db 149 LQDDPKETDNTSSDK 164

RESULT 7
GPSI_RAT
ID GPSI_RAT STANDARD; PRT; 471 AA.
AC P97834;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE G PROTEIN PATHWAY SUPPRESSOR 1 (GPSI PROTEIN) (MPH PROTEIN).
GN GPSI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=WISTAR; TISSUE=PINEAL GLAND;
RX MEDLINE; 99079337.
RA Nitz B., Ivell R., Hartung S., Middendorff R., Olcese J.;
RT "A fusca gene homologous in mammalian tissues: Developmental
regulation in the rat testes."
RL Mol. Cell. Endocrinol. 144:131-138(1998).
CC -|- FUNCTION: SUPPRESSES G-PROTEIN- AND MITOGEN-ACTIVATED PROTEIN
KINASE MEDIATED SIGNAL TRANSDUCTION.
CC -|- SIMILARITY: BELONGS TO THE FUS6 FAMILY.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X87885; CAA61139.1; -;
DR INTERPRO; IPR000717; -;
DR PFAM; PF01399; PCI; 1.
SQ SEQUENCE 471 AA; 53428 MW; ED919C9A9F3C8711 CRC64;

Query Match 41.6%; Score 42; DB 1; Length 471;
Best Local Similarity 37.5%; Pred. No. 17;
Matches 9; Conservative 5; Mismatches 4; Indels 6; Gaps 1;

QY 3 QDDPQNS-----IEATSVDSKQ 20
:||||: :|:|
Db 8 QEDPQNPADVNVVVENPTLDLEQ 31

RESULT 8
YD73_SCHPO
ID YD73_SCHPO STANDARD; PRT; 942 AA.
AC Q10328;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOPHETICAL 104.0 KDA PROTEIN C32A11.03C IN CHROMOSOME 1.
GN SPAC32A11.03C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
OC Schizosaccharomycetaceae; Schizosaccharomycetes.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -|- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -|- SIMILARITY: CONTAINS A HOMEBOX DOMAIN.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; Z69796; CAA93700.1; -;
DR HSP; P02836; 2HDD.
DR INTERPRO; IPR001356; -;
DR PFAM; PF00046; homeobox; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW Hypothetical protein; DNA-binding; Homeobox; Nuclear protein.
FT DNA_BIND 164 224 HOMEBOX.
SQ SEQUENCE 942 AA; 103986 MW; 3E5F22A7DA73F20F CRC64;

Query Match 41.1%; Score 41.5; DB 1; Length 942;
Best Local Similarity 40.9%; Pred. No. 45;
Matches 9; Conservative 6; Mismatches 2; Indels 5; Gaps 1;

QY 3 QDDPQNS-----IEATSVDSKQ 19
:||||: :||||:|
Db 126 EESPANSKQNEWEATSEKAK 147

RESULT 9
DCD_METH
ID DCD_METH STANDARD; PRT; 197 AA.
AC O27875;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROBABLE DEOXYCYTIDINE TRIPHOSPHATE DEAMINASE (EC 3.5.4.13) (DCTP
DE DEAMINASE).
GN DCD OR MTH1847.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RX MEDLINE; 98037514.

```
Query Match 45.5%; Score 46; DB 1; Length 612;
Best Local Similarity 42.1%; Pred. No. 4.8;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 IMQDDPQNSIATSVKSKQ 19
      ::::: | | | | |
DB 127 LLQDETENIDESTQVDQOQ 145

RESULT 4
GPS1_HUMAN STANDARD; PRT; 471 AA.
AC Q13098;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE G PROTEIN PATHWAY SUPPRESSOR 1 (GPS1 PROTEIN) (MFH PROTEIN).
GN GPS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97098647.
RA Spain B.H., Bowditch K.S., Pacal A., Fluckiger Staub S., Koo D.,
RA Chang K.-Y.R., Xie W., Colicelli J.;
RT "Two human cDNAs, including a homolog of Arabidopsis FUS6 (COP1),
RT suppress G-protein- and mitogen-activated protein kinase-mediated
RT signal transduction in yeast and mammalian cells."
RL Mol. Cell. Biol. 16:6698-6706(1996).
CC -1- FUNCTION: SUPPRESSES G-PROTEIN- AND MITOGEN-ACTIVATED PROTEIN
CC KINASE-MEDIATED SIGNAL TRANSDUCTION.
CC -1- SIMILARITY: BELONGS TO THE FUS6 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U20285; AAC50906.1; ALT_INIT.
DR MIN; 601934.
DR INTERPRO; IPR000717;
DR PFM; PFM01399; PCI; 1.
SQ SEQUENCE 471 AA; 53371 MW; 692E05F06BF7C0A3 CRC64;

Query Match 44.6%; Score 45; DB 1; Length 471;
Best Local Similarity 41.7%; Pred. No. 5.2;
Matches 10; Conservative 4; Mismatches 4; Indels 6; Gaps 1;

QY 3 QDDPQNS-----IEATSVKSKQ 20
      |::|||: | | | | |
DB 8 QEDQNPADVNVVVENPSLDLEQY 31

RESULT 5
YUFN_BACSU STANDARD; PRT; 350 AA.
AC O05252;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE HYPOTHETICAL LIPOPROTEIN YUFN PRECURSOR.
GN YUFN.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE FROM N.A.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U20285; AAC50906.1; ALT_INIT.
DR MIN; 601934.
DR INTERPRO; IPR000717;
DR PFM; PFM01399; PCI; 1.
SQ SEQUENCE 471 AA; 53371 MW; 692E05F06BF7C0A3 CRC64;

Query Match 43.6%; Score 44; DB 1; Length 350;
Best Local Similarity 35.0%; Pred. No. 5.4;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 IMQDDPQNSIATSVKSKQ 20
      :::|||: | | | | |
DB 236 LKEDPKRDVWVGVDKQY 255

RESULT 6
LHX5_MOUSE STANDARD; PRT; 402 AA.
AC P50459;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE LIM/HOMEOBOX PROTEIN LHX5 (HOMEOBOX PROTEIN LIM-2).
GN LHX5 OR LIM2 OR LIM-2.
OS Mus musculus (Mouse), and Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX SPECIES=RAT; TISSUE=BRAIN;
RX MEDLINE; 95094281.
RA Tsuchida T., Ensigni M., Morton S.B., Baldassare M., Edlund T.,
RA Jessell T.M., Pfaff S.L.;
RT "Topographic organization of embryonic motor neurons defined by
RT expression of LIM homeobox genes.";
RL Cell 79:957-970(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX SPECIES=MOUSE; STRAIN=FVB/N;
RX MEDLINE; 96411680.
RA Bertuzzi S., Sheng H.Z., Copeland N.G., Gilbert D.J., Jenkins N.A.,
RA Talra M., Dawid I.B., Westphal H.;
RT "Molecular cloning, structure, and chromosomal localization of the
RT mouse LIM/homeobox gene Lhx5."
RL Genomics 36:234-239(1996).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -1- SIMILARITY: TO OTHER HOMEOBOX DOMAINS. BELONGS TO THE LIM
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
CC IONS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Z93937; CAB07936.1;
DR EMBL; Z99120; CAB15143.1;
DR SUBTILIS; BG12349; YUFN.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Hypothetical protein; Membrane; Lipoprotein; Signal.
FT SIGNAL 1 13 PROBABLE.
FT CHAIN 14 350 HYPOTHETICAL LIPOPROTEIN YUFN.
FT LIPID 14 14 N-ACYL-DIGLYCERIDE (PROBABLE).
SQ SEQUENCE 350 AA; 37349 MW; 16D5176A52A99284 CRC64;

Query Match 43.6%; Score 44; DB 1; Length 350;
Best Local Similarity 35.0%; Pred. No. 5.4;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 IMQDDPQNSIATSVKSKQ 20
      :::|||: | | | | |
DB 236 LKEDPKRDVWVGVDKQY 255
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 09:03:45 ; Search time 58.45 seconds
(without alignments)
10.932 Million cell updates/sec

Title: US-09-236-468A-2_COPY_214_233
Perfect score: 101
Sequence: 1 IMQDDPQNSIATSVDSQY 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 87993

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB ID	Description
1	101	100.0	550	1	PTR2_HUMAN	P49190 homo sapien
2	60	59.4	546	1	PTH2_RAT	P70555 rattus norv
3	46	45.5	612	1	VBL_HPV1A	P03111 human papil
4	45	44.6	471	1	GSL1_HUMAN	Q13098 homo sapien
5	44	43.6	350	1	YUFN_BACSU	O05252 bacillus su
6	44	43.6	402	1	LHX5_MOUSE	P30459 mus musculu
7	42	41.6	471	1	GPS1_RAT	P97834 rattus norv
8	41.5	41.1	942	1	YD73_SCHPO	Q10328 schizosacch
9	41	40.6	197	1	DGD_METTH	O27875 methanobact
10	41	40.6	724	1	PEDD_PEDAC	P36497 pediococcus
11	41	40.6	730	1	GLN3_YEAST	P18494 saccharomyc
12	41	40.6	776	1	ISP4_SCHPO	P40900 schizosacch
13	40	39.6	238	1	EG_GOSHI	Q01197 gossypium h
14	40	39.6	255	1	143B_TOBAC	O49995 nicotiana t
15	40	39.6	256	1	OBFI_MOUSE	O64693 mus musculu
16	40	39.6	309	1	YL11_CAEEL	Q11098 caenorhabdi
17	40	39.6	580	1	Y586_METJA	Q58099 methanococc
18	40	39.6	611	1	LPL1_YEAST	P32487 saccharomyc
19	40	39.6	754	1	RAD4_YEAST	P14736 saccharomyc
20	40	39.6	1902	1	PLP_LACLC	P16271 lactococcus
21	40	39.6	1902	1	P2P_LACPA	Q02470 lactobacill
22	40	39.6	1902	1	P3P_LACLC	P15292 lactococcus
23	39.5	39.1	163	1	YA9F_SCHPO	Q09791 schizosacch
24	39.5	39.1	329	1	SRA4_CAEEL	Q09206 caenorhabdi
25	39	38.6	233	1	SERA_ENTHI	P21138 entamoeba h
26	39	38.6	338	1	G3P_METFO	P19315 methanobact
27	39	38.6	395	1	RPPP_NDVA	P16073 newcastle d
28	39	38.6	471	1	AUR1_CANAL	Q13332 candida alb
29	39	38.6	513	1	KAA8_SCHPO	Q09792 schizosacch
30	39	38.6	620	1	KEFC_ECOLI	P03819 escherichia
31	39	38.6	929	1	YDM6_SCHPO	P87137 schizosacch
32	39	38.6	1093	1	SW14_YEAST	P25302 saccharomyc
33	39	38.6	1590	1	GCN2_YEAST	P15442 saccharomyc

ALIGNMENTS

RESULT 1	ID	PTR2_HUMAN	STANDARD;	PRT;	550 AA.
AC	P49190;				
DT	01-FEB-1996	(Rel. 33, Created)			
DT	01-FEB-1996	(Rel. 33, Last sequence update)			
DT	01-OCT-2000	(Rel. 40, Last annotation update)			
DE	PARATHYROID HORMONE RECEPTOR PRECURSOR (PTH2 RECEPTOR).				
GN	PTH2.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
RC	SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.				
RC	TISSUE=BRAIN;				
RX	MEDLINE; 95318121.				
RA	Usdin T.B., Gruber C., Bonner T.I.;				
RT	"Identification and functional expression of a receptor selectively recognizing parathyroid hormone, the PTH2 receptor.";				
RL	J. Biol. Chem. 270:15455-15458(1995).				
RN	[2]				
RP	SEQUENCE OF 26-40 AND 306-550 FROM N.A.				
RX	MEDLINE; 97079671.				
RA	Usdin T.B., Modi W., Bonner T.I.;				
RT	"Assignment of the human PTH2 receptor gene (PTH2) to chromosome 2q33 by fluorescence in situ hybridization.";				
RL	Genomics 37:140-141(1996)				
CC	-!- FUNCTION: THIS IS A SPECIFIC RECEPTOR FOR PARATHYROID HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYL CYCLASE.				
CC	-!- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN BRAIN AND PANCREAS.				
CC	ALSO EXPRESSED IN THE TESTIS.				
CC	-!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.				
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).				
CC	EMBL; U25128; AAC50157.1; -				
DR	EMBL; U47124; AAA96796.1; -				
DR	EMBL; U47129; AAC50767.1; -				
DR	EMBL; U47125; AAC50767.1; JOINED.				
DR	EMBL; U47126; AAC50767.1; JOINED.				
DR	EMBL; U47127; AAC50767.1; JOINED.				
DR	EMBL; U47128; AAC50767.1; JOINED.				
DR	GCRDB; GCR.2003; -				
DR	MIM; 601469; -				
DR	INTERPRO; IPR000832; -				
DR	PFAM; PF00002; 7tm_2; 1.				
DR	PRINTS; PR00249; GPCRSECRETIN.				
DR	PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.				
DR	PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.				

34	39	38.6	2493	1	YBA4_YEAST	P35194 saccharomyc
35	38.5	38.1	124	1	GLU1_LOPAM	P01278 lophius ahe
36	38	37.6	440	1	CHLB_CYAPA	P48099 cyanophora
37	38	37.6	471	1	RNT1_YEAST	Q02555 saccharomyc
38	38	37.6	473	1	TBG_YEAST	P53378 saccharomyc
39	38	37.6	499	1	PTFA_HAEIN	P44715 h pts syste
40	38	37.6	870	1	YCS0_YEAST	P25623 saccharomyc
41	38	37.6	957	1	NED4_MOUSE	P46935 mus musculu
42	38	37.6	1208	1	AXL1_YEAST	P40851 saccharomyc
43	38	37.6	1442	1	Y39F_YEAST	P47169 saccharomyc
44	38	37.6	1451	1	SPT6_YEAST	P23615 saccharomyc
45	38	37.6	2375	1	ATRX_HUMAN	P46100 homo sapien

A:Reference number: A55198; MUID:95094281
A:Accession: I61573
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-402 <RES>
A:Cross-references: GB:L35572; NID:g531219; PIDN:AAA62162.1; PID:g531220
C:Genetics:
A:Gene: lim-2
C:Superfamily: homeotic protein lim-1; homeobox homology; LIM metal-binding repeat homolog
C:Keywords: DNA binding; duplication; embryo; homeobox; nucleus; transcription regulation
F:5-55/Domain: LIM metal-binding repeat homology <LIM1>
F:64-118/Domain: LIM metal-binding repeat homology <LIM2>
F:181-237/Domain: homeobox homology <HOX>

Query Match 43.6%; Score 44; DB 1; Length 402;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 MQDDPQNSIEATSVDK 17
:||||: :|||
Db 149 LQDDPKETDNTSSDK 164

RESULT 8
BB1923
hypothetical protein NMA0785 [imported] - Neisseria meningitidis (group A strain Z2491)
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000
C:Accession: B81923
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
: Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556
A:Accession: B81923
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-149 <PAR>
A:Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84068.1; PID:g737950
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA0785

Query Match 42.6%; Score 43; DB 2; Length 149;
Best Local Similarity 42.1%; Pred. No. 6.6;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 2 MQDDPQNSIEATSVDKSOY 20
||| | :|||
Db 106 MQSDLQKKLNASYNKQY 124

RESULT 9
F75112
hypothetical protein PAB1900 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: F75112
R:Anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: F75112
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-394 <KAW>
A:Cross-references: GB:AJ248285; GB:AL096836; NID:g5458067; PIDN:CAB49615.1; PID:e151551
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB1900

Query Match 41.6%; Score 42; DB 2; Length 618;
Best Local Similarity 37.5%; Pred. No. 49;
Matches 6; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 1 IMQDDPQNSIEATSV 16
::|:| |::|||
Db 534 LLLDEPTNHLDTISID 549

RESULT 12
T38649
hypothetical homeobox domain protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T38649
R:Skelton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

Query Match 42.6%; Score 43; DB 2; Length 394;
Best Local Similarity 42.9%; Pred. No. 20;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 2 MQDDPQNSIEATSV 15
| ||| : |||
Db 262 MSDDPEENLKATTI 275

RESULT 10
S57449
fusca protein homolog - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 05-Nov-1999
C:Accession: S57449
R:Nitz, B.; Hartung, S.; Ivell, R.; Olcese, J.
submitted to the EMBL Data Library, June 1995
A:Reference number: S57449
A:Accession: S57449
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-471 <NIT>
A:Cross-references: EMBL:X87885; NID:g871527; PIDN:CAA61139.1; PID:g871528

Query Match 41.6%; Score 42; DB 2; Length 471;
Best Local Similarity 37.5%; Pred. No. 36;
Matches 9; Conservative 5; Mismatches 4; Indels 6; Gaps 1;

Qy 3 QDDPQNS-----IEATSVDKSQY 20
|:||||: :| :|
Db 8 QEDPQNPDPVNVVENPTLDLEQY 31

RESULT 11
T39617
probable ABC transporter - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: T39617
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.
submitted to the EMBL Data Library, March 1998
A:Reference number: Z21843
A:Accession: T39617
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-618 <WOO>
A:Cross-references: EMBL:AL022104; PIDN:CAA17906.1; GSPDB:GN00067; SPDB:SPBC16H5.08c
A:Experimental source: strain 972h-; cosmid c16H5
C:Genetics:
A:Gene: SPDB:SPBC16H5.08c
A:Map position: 2
A:Introns: 12/3
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog

C:Accession: G01646
R:Bowdish, K.S.
submitted to the EMBL Data Library, January 1995
A:Reference number: G08025
A:Accession: G01646
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-500 <HOW>
A:Cross-references: EMBL:U20285; NID:g644878; PIDN:AAC50906.1; PID:g644879
C:Genetics:
A:Gene: GPS1

Query Match 44.6%; Score 45; DB 2; Length 500;
Best Local Similarity 41.7%; Pred. No. 12;
Matches 10; Conservative 4; Mismatches 4; Indels 6; Gaps 1;

QY 3 ODDPONS-----LEATSVDKSOY 20
:||||: :| ||| ||
Db 37 QEDPONAPDWNVVVNPSLLEQY 60

RESULT 6
C70009
ABC transporter (lipoprotein) homolog yufN - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: C70009
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.: Bron, S.; Brouillet, S.; Bruschii, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.: Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.;
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portecel
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil
A:Reference number: A69580; MUID:98044033
A:Accession: C70009
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-350 <GUN>
A:Cross-references: GB:Z99120; GB:AL009126; NID:g2635613; PID:g2635650
A:Experimental source: strain 168
C:Genetics:
A:Gene: yufN
C:Superfamily: ABC transporter yufN

Query Match 43.6%; Score 44; DB 1; Length 350;
Best Local Similarity 35.0%; Pred. No. 12;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 IMQDDPONSLEATSVDKSOY 20
:||||: : ||| ||
Db 236 LKKEDPRDWNVIGVDRDQY 255

RESULT 7
I61573
homeotic protein lim-2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
C:Accession: I61573
R:Tsuchida, T.; Ensigni, M.; Morton, S.B.; Baldassare, M.; Edlund, T.; Jessell, T.M.;
Cell 79, 957-970, 1994
A:Title: Topographic organization of embryonic motor neurons defined by expression of

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: November 8, 2000, 08:53:24 ; Search time 99.87 Seconds
(without alignments)
12.709 Million cell updates/sec

Title: US-09-236-468a-2_COPY_214_233
Perfect score: 101
Sequence: 1 IMQDDPQNSIEATSVDKSQY 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues
Total number of hits satisfying chosen parameters: 182106

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	550	2 A57519	parathyroid hormon
2	46	45.5	612	1 W1WLE	E1 protein - human
3	46	45.5	1876	2 T13351	phosphoinositide 3
4	46	45.5	1876	2 T13801	phosphoinositide 3
5	45	44.6	500	2 G01646	fusca protein homo
6	44	43.6	350	1 C70009	ABC transporter (1
7	44	43.6	402	1 I61573	homeotic protein 1
8	43	42.6	149	2 B81923	hypothetical prote
9	43	42.6	394	2 F75112	hypothetical prote
10	42	41.6	471	2 S57449	fusca protein homo
11	42	41.6	618	2 T39617	probable ABC trans
12	41.5	41.1	942	2 T38649	hypothetical homeo
13	41	40.6	200	2 A69114	dCTP deaminase (EC
14	41	40.6	207	2 D81808	hypothetical prote
15	41	40.6	211	2 S69932	outer surface prot
16	41	40.6	724	2 D48941	ATP-dependent tran
17	41	40.6	730	2 S50543	GLN3 protein - yea
18	41	40.6	776	2 S45495	isp4 protein - fis
19	41	40.6	905	2 T38314	probable vacuolar
20	41	40.6	940	2 T43263	cell division prot
21	41	40.6	920	2 T50433	septum positioning
22	40.5	40.1	887	2 T70642	ubiquitin ligase N
23	40	39.6	66	2 C72028	conserved hypothet
24	40	39.6	238	2 A46130	fiber protein E6 (
25	40	39.6	241	2 S65062	fiber protein E6 (
26	40	39.6	246	2 S65063	fiber protein E6 (
27	40	39.6	255	2 T04127	14-3-3 protein, is
28	40	39.6	256	2 S63588	transcription coac
29	40	39.6	268	2 S71830	transcription coac

30	40	39.6	309	2 T34080	hypothetical prote
31	40	39.6	325	2 T25122	hypothetical prote
32	40	39.6	355	2 D81923	hypothetical prote
33	40	39.6	580	2 F64385	hypothetical prote
34	40	39.6	609	2 T16135	hypothetical prote
35	40	39.6	611	2 S60914	lysine transport p
36	40	39.6	754	1 DDBYD4	excision repair pr
37	40	39.6	1130	2 T20288	hypothetical prote
38	40	39.6	1902	1 B44858	lactocepin (EC 3.4
39	40	39.6	1902	2 B45764	lactocepin (EC 3.4
40	40	39.6	1962	2 A32634	lactocepin (EC 3.4
41	39.5	39.1	163	2 S62444	conserved hypothet
42	39.5	39.1	329	2 T18615	hypothetical prote
43	39	38.6	251	2 T34622	NADH dehydrogenase
44	39	38.6	291	2 S75791	hypothetical prote
45	39	38.6	338	1 S02803	glyceraldehyde-3-p

ALIGNMENTS

RESULT 1
A57519
parathyroid hormone receptor 2 precursor - human
N: Alternate names: PTH2 receptor
C: Species: Homo sapiens (man)
C: Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 24-Nov-1999
C: Accession: A57519
R: Usdin, T.B.; Gruber, C.; Bonner, T.I.
J. Biol. Chem. 270, 15455-15458, 1995
A: Title: Identification and functional expression of a receptor selectively recognizing
A: Reference number: A57519; MUID:95318121
A: Accession: A57519
A: Status: preliminary; nucleic acid sequence not shown
A: Molecule type: mRNA
A: Residues: 1-550 <USD>
A: Cross-references: GB:U25128; NID:9887966; PIDN:RAC50157.1; PID:9887967
C: Genetics:
A: Gene: GDB:PTHR2; PTHR2R
A: Cross-references: GDB:731977; OMIM:601469
A: Map position: 2q33-2q33
C: Superfamily: glucagon receptor
C: Keywords: hormone receptor

Query Match 100.0%; Score 101; DB 2; Length 550;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IMQDDPQNSIEATSVDKSQY 20
|||||
DB 214 IMQDDPQNSIEATSVDKSQY 233

RESULT 2
W1WLE
E1 protein - human papillomavirus type 1a
C: Species: human papillomavirus type 1a
C: Date: 18-Aug-1982 #sequence_revision 14-Nov-1983 #text_change 24-Feb-1994
C: Accession: A03655
R: Danos, O.; Engel, L.W.; Chen, E.Y.; Yaniv, M.; Howley, P.M.
J. Virol. 46, 557-566, 1983
A: Title: Comparative analysis of the human type 1a and bovine type 1 papillomavirus g
A: Reference number: A92993; MUID:83189357
A: Accession: A03655
A: Molecule type: DNA
A: Residues: 1-612 <DAN>
A: Note: the authors translated the codon TCA for residue 274 as Arg, CCT for residue
C: Superfamily: papillomavirus E1 protein
C: Keywords: early protein
Query Match 45.5%; Score 46; DB 1; Length 612;

THIS PAGE BLANK (USPTO)

```

; SOFTWARE: DOS Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/075,545
; FILING DATE: 14-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP91/01722
; FILING DATE: 17-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fleit, Martin; Gollin, Michael A.
; REGISTRATION NUMBER: 16,900; 31,957
; REFERENCE/DOCKET NUMBER: 47004-015
; TELEPHONE: (202) 789-3400
; TELEFAX: (202) 789-1158
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-075-545-2

```

```

Query Match 36.6%; Score 37; DB 1; Length 126;
Best Local Similarity 36.8%; Pred. No. 26;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy 1 IMQDDPQNSIEATSVDSQ 19
   :: || || || ::
Db 68 VVELDPSAKIEVYYDRNK 86

```

Search completed: November 8, 2000, 08:49:18
Job time: 111 sec

NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4103US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 50
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-178-477B-14

Query Match 36.6%; Score 37; DB 1; Length 50;
Best Local Similarity 42.1%; Pred. No. 8.4;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 2 MQDDPONSIEATSVDKSQY 30
Db 18 VQDVKSGSIQSSDDKWQF 36

RESULT 13
US-08-256-261-10
Sequence 10, Application US/08256261
Patent No. 5801037

GENERAL INFORMATION:
APPLICANT: Behnke, Detlef
APPLICANT: Schlott, Bernhard
APPLICANT: Albrecht, Sybille
APPLICANT: G hrs, Karl-Heinz
APPLICANT: Hartmann, Manfred
TITLE OF INVENTION: Expression of signal-peptide-free
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave

STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/256,261
FILING DATE:

CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-256-261-10

Query Match 36.6%; Score 37; DB 1; Length 123;
Best Local Similarity 36.8%; Pred. No. 25;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 IMQDDPONSIEATSVDKSQY 19
Db 65 VVELDPSAKIEVTYYDKNK 83

RESULT 14
US-08-852-299-10
Sequence 10, Application US/08852299
Patent No. 6010897
GENERAL INFORMATION:
APPLICANT: Behnke, Detlef
APPLICANT: Schlott, Bernhard
APPLICANT: Albrecht, Sybille
APPLICANT: G hrs, Karl-Heinz
APPLICANT: Hartmann, Manfred
TITLE OF INVENTION: Expression of signal-peptide-free
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,299
FILING DATE: 17-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/256,261
FILING DATE:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-852-299-10

Query Match 36.6%; Score 37; DB 3; Length 123;
Best Local Similarity 36.8%; Pred. No. 25;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 IMQDDPONSIEATSVDKSQY 19
Db 65 VVELDPSAKIEVTYYDKNK 83

RESULT 15
US-08-075-545-2
Sequence 2, Application US/08075545
Patent No. 5475089

GENERAL INFORMATION:
APPLICANT: MATSUO, Osamu; SAKAI, Masashi; SHIMURA,
APPLICANT: KISAKU; SANSABA, Hiroshi; WATANABE,
APPLICANT: TSUNEKAZU; HASHIMOTO, Tsuneo; SHISHIDO,
APPLICANT: YOSHIYUKI; HASHIMOTO, Shusuke; YOKOKURA,
APPLICANT: TERUO; ONOUE, Masaharu; SAKO, Tomoyuki
TITLE OF INVENTION: THROMBOLYTIC AGENT
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: KECK, MAHIN & CATE
STREET: P.O. BOX 06110
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: U.S.A.
ZIP: 60606-0110

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS/PC DOS

```

; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Evans, Cheryl
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/686,878A
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 75 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-686-878A-12

```

```

Query Match 37.6%; Score 38; DB 1; Length 75;
Best Local Similarity 46.7%; Pred. No. 9.3;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy 1 IMQDDPNSIEATSV 15
; : | | | | : | | |
Db 21 VFPDKPNSDKATSL 35

```

```

RESULT 11
US-08-447-031A-2
; Sequence 2, Application US/08447031A
; Patent No. 5851794
; GENERAL INFORMATION:
; APPLICANT: GUSS, Bengt
; APPLICANT: HOOK, Magnus
; APPLICANT: JONSSON, Hans
; APPLICANT: LINDBERG, Martin
; APPLICANT: PATTI, Joseph
; APPLICANT: SIGNAS, Christer
; APPLICANT: SWITALSKI, Lech
; TITLE OF INVENTION: A COLLAGEN BINDING PROTEIN AS WELL AS
; TITLE OF INVENTION: ITS PREPARATION
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,031A
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/861,804
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/SE91/00707
; FILING DATE: 22-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9003374-7
; FILING DATE: 22-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 012889-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1183 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-447-031A-2

```

```

Query Match 37.1%; Score 37.5; DB 2; Length 1183;
Best Local Similarity 52.9%; Pred. No. 3.2e+02;
Matches 9; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

```

```

Qy 3 QDDPQNSIEATSVDKSQ 19
; | | | | | | | | | |
Db 458 QDDNQNT---TPVDKAE 471

```

```

RESULT 12
US-08-178-477B-14
; Sequence 14, Application US/08178477B
; Patent No. 5756343
; GENERAL INFORMATION:
; APPLICANT: WU, CARL; CLOS, JOACHIM;
; APPLICANT: WESTWOOD, J. TIMOTHY.; RABINDRAN, SRIDHAR
; TITLE OF INVENTION: CELL STRESS
; TITLE OF INVENTION: TRANSCRIPTIONAL FACTORS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/178,477B
; FILING DATE: 07-JAN-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/617,910
; FILING DATE: 26-NOV-1990
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:

```

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/243,403
;; FILING DATE: May 16, 1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Warren M. Cheek, Jr.
;; REGISTRATION NUMBER: 33,367
;; REFERENCE/DOCKET NUMBER:
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-371-8850
;; TELEFAX:
;; TELEX:

;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 471
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-423-752-14

Query Match 38.6%; Score 39; DB 3; Length 471;
Best Local Similarity 45.0%; Pred. No. 58;
Matches 9; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

Qy 1 IMQDDPQNSI----EATSVD 16
; :||| |
Db 445 VFEDEPGSTVAASSATSVD 464

RESULT 8

US-08-492-459-22
; Sequence 22, Application US/08492459
; Patent No. 6015689
; GENERAL INFORMATION:
; APPLICANT: Takashi OKADO et al.

;; TITLE OF INVENTION: REGULATION OF AUREOBASIDIN SENSITIVITY IN FUNGUS
;; NUMBER OF SEQUENCES: 36
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Wenderoth, Lind & Ponack
;; STREET: 805 Fifteenth Street, N.W., #700
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: U.S.A.
;; ZIP: 20005

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 mb
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: MS-DOS
;; SOFTWARE: Wordperfect 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/492,459
;; FILING DATE: June 20, 1995
;; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/243,403
;; FILING DATE: May 16, 1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Warren M. Cheek, Jr.
;; REGISTRATION NUMBER: 33,367
;; REFERENCE/DOCKET NUMBER:
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-371-8850
;; TELEFAX:
;; TELEX:

;; INFORMATION FOR SEQ ID NO: 22:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 472
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-492-459-22

Query Match 38.6%; Score 39; DB 3; Length 471;
Best Local Similarity 45.0%; Pred. No. 58;
Matches 9; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

Qy 1 IMQDDPQNSI----EATSVD 16
; :||| |
Db 445 VFEDEPGSTVAASSATSVD 464

Query Match 38.6%; Score 39; DB 3; Length 472;
Best Local Similarity 45.0%; Pred. No. 59;
Matches 9; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

Qy 1 IMQDDPQNSI----EATSVD 16
; :||| |
Db 446 VFEDEPGSTVAASSATSVD 465

RESULT 9

US-08-423-752-22
; Sequence 22, Application US/08423752
; Patent No. 6022949
; GENERAL INFORMATION:
; APPLICANT: Takashi OKADO et al.

;; TITLE OF INVENTION: A GENE CODING FOR A PROTEIN REGULATING
;; TITLE OF INVENTION: AUREOBASIDIN SENSITIVITY
;; NUMBER OF SEQUENCES: 22
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Wenderoth, Lind & Ponack
;; STREET: 805 Fifteenth Street, N.W., #700
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: U.S.A.
;; ZIP: 20005

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: MS-DOS
;; SOFTWARE: Wordperfect 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/423,752
;; FILING DATE: April 18, 1995
;; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/243,403
;; FILING DATE: May 16, 1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Warren M. Cheek, Jr.
;; REGISTRATION NUMBER: 33,367
;; REFERENCE/DOCKET NUMBER:
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-371-8850
;; TELEFAX:
;; TELEX:

;; INFORMATION FOR SEQ ID NO: 22:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 472
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-423-752-22

Query Match 38.6%; Score 39; DB 3; Length 472;
Best Local Similarity 45.0%; Pred. No. 59;
Matches 9; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

Qy 1 IMQDDPQNSI----EATSVD 16
; :||| |
Db 446 VFEDEPGSTVAASSATSVD 465

RESULT 10

US-08-686-878A-12
; Sequence 12, Application US/08686878A
; Patent No. 5708157
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John

;; TITLE OF INVENTION: A GENE CODING FOR A PROTEIN REGULATING
;; TITLE OF INVENTION: AUREOBASIDIN SENSITIVITY
;; NUMBER OF SEQUENCES: 22
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Wenderoth, Lind & Ponack
;; STREET: 805 Fifteenth Street, N.W., #700
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: U.S.A.
;; ZIP: 20005

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: MS-DOS
;; SOFTWARE: Wordperfect 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/423,752
;; FILING DATE: April 18, 1995
;; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/243,403
;; FILING DATE: May 16, 1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Warren M. Cheek, Jr.
;; REGISTRATION NUMBER: 33,367
;; REFERENCE/DOCKET NUMBER:
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-371-8850
;; TELEFAX:
;; TELEX:

;; INFORMATION FOR SEQ ID NO: 22:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 472
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-423-752-22

Query Match 38.6%; Score 39; DB 3; Length 472;
Best Local Similarity 45.0%; Pred. No. 59;
Matches 9; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

Qy 1 IMQDDPQNSI----EATSVD 16
; :||| |
Db 446 VFEDEPGSTVAASSATSVD 465

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-609-049A-12

Query Match 45.5%; Score 46; DB 2; Length 1876;
Best Local Similarity 56.2%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 MODDPQNSIEATSVDK 17
Db 1187 LPDDPHNSIGAAMVDQ 1202

RESULT 5
US-08-609-049A-28
; Sequence 28, Application US/08609049A
; Patent No. 5948664
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Men
; TITLE OF INVENTION: No. 5948664el PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,049A
; FILING DATE: 29-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-0637000S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1876 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-609-049A-28

Query Match 45.5%; Score 46; DB 2; Length 1876;
Best Local Similarity 56.2%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 MODDPQNSIEATSVDK 17
Db 1187 LPDDPHNSIGAAMVDQ 1202

RESULT 6
US-08-492-459-14
; Sequence 14, Application US/08492459
; Patent No. 6015689
; GENERAL INFORMATION:
; APPLICANT: Takashi OKADO et al.

; TITLE OF INVENTION: REGULATION OF AUROBASIDIN SENSITIVITY IN FUNGUS
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/492,459
; FILING DATE: June 20, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/243,403
; FILING DATE: May 16, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 471
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-492-459-14

Query Match 38.6%; Score 39; DB 3; Length 471;
Best Local Similarity 45.0%; Pred. No. 58;
Matches 9; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

Qy 1 IMWDDPQNSI---EATSVSD 16
Db 445 VFEDEPQGSTYAASSATSVSD 464

RESULT 7
US-08-423-752-14
; Sequence 14, Application US/08423752
; Patent No. 6022949
; GENERAL INFORMATION:
; APPLICANT: Takashi OKADO et al.
; TITLE OF INVENTION: A GENE CODING FOR A PROTEIN REGULATING
; TITLE OF INVENTION: AUROBASIDIN SENSITIVITY
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/423,752
; FILING DATE: April 18, 1995
; CLASSIFICATION: 435

Db 1 IMODDPQNSIEATSVDKSQY 20

RESULT 2

US-08-468-011A-2
; Sequence 2, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Steven
; APPLICANT: Ruben, Craig A
; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTDG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-011A-2

Query Match 100.0%; Score 101; DB 3; Length 541;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IMODDPQNSIEATSVDKSQY 20

Db 214 IMODDPQNSIEATSVDKSQY 233

RESULT 3

PCT-US95-07085-2
; Sequence 2, Application PC/TUS9507085
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone Receptor
; TITLE OF INVENTION: HLTDG74
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland

; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07085
; FILING DATE: 05-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-393
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-07085-2

Query Match 100.0%; Score 101; DB 4; Length 541;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IMODDPQNSIEATSVDKSQY 20

Db 214 IMODDPQNSIEATSVDKSQY 233

RESULT 4

US-08-609-049A-12
; Sequence 12, Application US/08609049A
; Patent No. 5948664
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 5948664el PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,049A
; FILING DATE: 29-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1876 amino acids

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:49:17 ; Search time 97.15 Seconds
(without alignments)
3.451 Million cell updates/sec

Title: US-09-236-468A-2_COPY_214_233

Perfect score: 101
Sequence: 1 IMQDDPQNSIEATSDKSOY 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 164575

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	101	100.0	60	3	US-08-468-011A-9
2	101	100.0	541	3	US-08-468-011A-2
3	101	100.0	541	4	PCT-US95-07085-2
4	46	45.5	1876	2	US-08-609-049A-12
5	46	45.5	1876	2	US-08-609-049A-28
6	39	38.6	471	3	US-08-492-459-14
7	39	38.6	471	3	US-08-423-752-14
8	39	38.6	472	3	US-08-492-459-22
9	39	38.6	472	3	US-08-423-752-22
10	38	37.6	75	1	US-08-686-878A-12
11	37.5	37.1	1183	2	US-08-447-031A-2
12	37	36.6	50	1	US-08-178-477B-14
13	37	36.6	123	1	US-08-256-261-10
14	37	36.6	123	3	US-08-852-299-10
15	37	36.6	126	1	US-08-075-545-2
16	37	36.6	127	1	US-08-256-261-8
17	37	36.6	127	3	US-08-852-299-8
18	37	36.6	136	1	US-08-371-505-2
19	37	36.6	136	2	US-08-784-971-5
20	37	36.6	137	1	US-08-256-261-2
21	37	36.6	137	1	US-08-256-261-4
22	37	36.6	137	1	US-08-256-261-6
23	37	36.6	137	1	US-08-256-261-12
24	37	36.6	137	1	US-08-256-261-14
25	37	36.6	137	3	US-08-852-299-2
26	37	36.6	137	3	US-08-852-299-4
27	37	36.6	137	3	US-08-852-299-6
28	37	36.6	137	3	US-08-852-299-12

29	37	36.6	137	3	US-08-852-299-14	Sequence 14, Appl
30	37	36.6	163	1	US-08-075-545-1	Sequence 1, Appl
31	37	36.6	163	1	US-08-256-261-17	Sequence 17, Appl
32	37	36.6	163	3	US-08-852-299-17	Sequence 17, Appl
33	37	36.6	225	1	US-08-462-169B-22	Sequence 22, Appl
34	37	36.6	225	2	US-08-951-822-25	Sequence 22, Appl
35	37	36.6	225	3	US-09-103-079-22	Sequence 22, Appl
36	37	36.6	225	3	US-08-705-245-10	Sequence 10, Appl
37	37	36.6	480	1	US-07-803-636A-2	Sequence 2, Appl
38	37	36.6	630	3	US-08-860-091A-2	Sequence 2, Appl
39	36.5	36.1	663	1	US-08-765-081-7	Sequence 7, Appl
40	36.5	36.1	663	3	US-09-098-082-7	Sequence 7, Appl
41	36.5	36.1	703	4	PCT-US95-06994-9	Sequence 9, Appl
42	36	35.6	27	1	US-07-971-096-6	Sequence 6, Appl
43	36	35.6	27	1	US-08-175-096-6	Sequence 6, Appl
44	36	35.6	423	1	US-08-190-802A-61	Sequence 61, Appl
45	36	35.6	535	3	US-08-286-870A-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-468-011A-9
; Sequence 9, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468.011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-011A-9

Query Match 100.0%; Score 101; DB 3; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.2e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IMQDDPQNSIEATSDKSOY 20
|||||

2

RESULT 13

R10941
ID R10941 standard; Protein; 1968 AA.

XX AC R10941;

XX DT 15-APR-1991 (first entry)

XX DE Mutant protease (delta137-139/ins9).

XX KW Mutant protease gene; fermentation; foodstuff; flavouring;
lactic acid bacteria.

XX OS Lactococcus lactis SK11.

XX FH Key Location/Qualifiers
Peptide 1..187

XX FT /label= sig_peptide

XX FT 188..1968

XX FT /label= mat_protein

XX PN EP411715-A.

XX PD 06-FEB-1991.

XX PF 02-AUG-1990; 90EP-0202113.

XX PR 04-AUG-1989; 89NL-0002010.

XX PA (NEZU-) NED INST ZUIVELONDE.

XX PI Vos PAJ, Siezen RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ;

XX DR WPI; 1991-038622/06.

XX DR N-PSDB; Q10871.

XX PT Mutant protease gene(s) and protease(s) - derived from type I and
III protease genes from lactococcal strains, used in fermentation
foodstuffs and flavourings

XX PS Disclosure; Fig 1(1-7)+5(b); 29pp; English.

XX CC The wild-type L.lactis SK11 protease gene sequence was determined
by the applicant (EP-307011).

XX CC The mutant protease having new cleavage specificities is

XX CC obtained by deleting three amino acids (nine bps) and inserting

XX CC 9 other residues.

XX CC This mutant may then be used to prepare hybrid proteases,

XX CC the fusion being between a type I and a type III protease of

XX CC L.lactis Wg2 and SK11.

XX CC The product has modified properties, e.g. thermostability,

XX CC alkaline/acid pH stability, oxidative stability, autolysis etc.,

XX CC compared to the parent protease(s). The proteases can be used for

XX CC preparing products (butter cheese, human and animal foodstuffs)

XX CC prepared with the aid of lactic acid bacteria.

XX CC See also Q10411-17 and Q10870-71.

XX SQ Sequence 1968 AA;

Query Match

Best Local Similarity 39.6%; Score 40; DB 12; Length 1968;

Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 IMQDDPONSIEATSVKSOY 20

: :||| | | | | | | | |

Db 1420 vytndpnfqtgtatdnag 1439

RESULT 14

R10940

ID R10940 standard; Protein; 1974 AA.

XX

AC R10940;

XX DT 15-APR-1991 (first entry)

XX DE Mutant protease (delta137-139/ins15).

XX KW Mutant protease gene; fermentation; foodstuff; flavouring;
lactic acid bacteria.

XX OS Lactococcus lactis SK11.

XX FH Key Location/Qualifiers
Peptide 1..187

XX FT /label= SIG_PEPTIDE

XX FT 188..1974

XX FT /label= MAT_PROTEIN

XX PN EP411715-A.

XX PD 06-FEB-1991.

XX PF 02-AUG-1990; 90EP-0202113.

XX PR 04-AUG-1989; 89NL-0002010.

XX PA (NEZU-) NED INST ZUIVELONDE.

XX PI Vos PAJ, Siezen RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ;

XX DR WPI; 1991-038622/06.

XX DR N-PSDB; Q10870.

XX PT Mutant protease gene(s) and protease(s) - derived from type I and
III protease genes from lactococcal strains, used in fermentation
foodstuffs and flavourings

XX PS Disclosure; Fig 1(1-7)+5(b); 29pp; English.

XX CC The wild-type L.lactis SK11 protease gene sequence was determined
by the applicant (EP-307011).

XX CC The mutant protease having new cleavage specificities is

XX CC obtained by deleting three amino acids (nine bps) and inserting

XX CC 15 other residues.

XX CC This mutant may then be used to prepare hybrid proteases,

XX CC the fusion being between a type I and a type III protease of

XX CC L.lactis Wg2 and SK11.

XX CC The product has modified properties, e.g. thermostability,

XX CC alkaline/acid pH stability, oxidative stability, autolysis etc.,

XX CC compared to the parent protease(s). The proteases can be used for

XX CC preparing products (butter cheese, human and animal foodstuffs)

XX CC prepared with the aid of lactic acid bacteria.

XX CC See also Q10411-17 and Q10870-71.

XX SQ Sequence 1974 AA;

Query Match

Best Local Similarity 39.6%; Score 40; DB 12; Length 1974;

Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 IMQDDPONSIEATSVKSOY 20

: :||| | | | | | | | |

Db 1426 vytndpnfqtgtatdnag 1445

RESULT 15

R14361

ID R14361 standard; Protein; 233 AA.

XX R14361;

XX AC

XX DT 23-JAN-1992 (first entry)

XX

CC See also Q10411-17 and Q10870-71.

XX Sequence 1962 AA;

Query Match 39.6%; Score 40; DB 12; Length 1962;
Best Local Similarity 35.0%; Pred. No. 4.6e+02;
Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 IMODDPQNSIEATSVKSKY 20
: || | | : ||
Db 1414 vytndpnfqtgtatdnaqy 1433

RESULT 11

R10559 ID R10559 standard; Protein; 1962 AA.

XX AC R10559;

XX 15-APR-1991 (first entry)

XX Mutant protease (A137G/K138L/T139A).

XX Mutant protease gene; fermentation; foodstuff; flavouring;
KW lactic acid bacteria.

XX Lactococcus lactis SK11.

XX Key Location/Qualifiers
FT Peptide 1..187

FT /label= signal_peptide
FT 188..1962

FT /label= mature_protein

XX EP411715-A.

XX 06-FEB-1991.

XX 02-AUG-1990; 90EP-0202113.

XX 04-AUG-1989; 89NL-0002010.

XX (NEZU-) NED INST ZUIVELONDE.

XX Vos PAJ, Siezen RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ;

XX WPI; 1991-038622/06.

XX N-PSDB; Q10413.

XX Mutant protease gene(s) and protease(s) - derived from type I and
PT III protease genes from lactococcal strains, used in fermentation
PT foodstuffs and flavourings

XX Disclosure; Fig 1(1-7)+5(b); 29 pp; English.

XX The mutant protease having new cleavage specificities is
CC obtained by replacing three amino acids.

CC This mutant may then be used to prepare hybrid proteases, of
CC the fusion being between a type I and a type III protease of

CC L.lactis Wg2 and SK11.

CC The product has modified properties, e.g. thermostability,
CC alkaline/acid pH stability, oxidative stability, autolysis etc.,

CC compared to the parent protease(s). The proteases can be used for
CC preparing products (butter cheese, human and animal foodstuffs)

CC prepared with the aid of lactic acid bacteria.

CC See also Q10411-17 and Q10870-71.

XX Sequence 1962 AA;

XX Query Match 39.6%; Score 40; DB 12; Length 1962;

XX Best Local Similarity 35.0%; Pred. No. 4.6e+02;

XX Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 IMODDPQNSIEATSVKSKY 20

: || | | : ||

Db 1414 vytndpnfqtgtatdnaqy 1433

Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 IMODDPQNSIEATSVKSKY 20

: || | | : ||

Db 1414 vytndpnfqtgtatdnaqy 1433

RESULT 12

R10563 ID R10563 standard; Protein; 1962 AA.

XX AC R10563;

XX 15-APR-1991 (first entry)

XX Mutant protease (K748T).

XX Mutant protease gene; fermentation; foodstuff; flavouring;
KW lactic acid bacteria.

XX Lactococcus lactis SK11.

XX Key Location/Qualifiers
FT Peptide 1..187

FT /label= signal_peptide

FT 188..1962

FT /label= mature_protein

XX EP411715-A.

XX 06-FEB-1991.

XX 02-AUG-1990; 90EP-0202113.

XX 04-AUG-1989; 89NL-0002010.

XX (NEZU-) NED INST ZUIVELONDE.

XX Vos PAJ, Siezen RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ;

XX WPI; 1991-038622/06.

XX N-PSDB; Q10417.

XX Mutant protease gene(s) and protease(s) - derived from type I and

PT III protease genes from lactococcal strains, used in fermentation

PT foodstuffs and flavourings

XX Disclosure; Fig 1(1-7)+5(a); 29 pp; English.

XX The mutant protease K748T having new cleavage specificities is

CC obtained by carrying out single amino acid substitutions.

CC This mutant may then be used to prepare hybrid proteases, of

CC the fusion being between a type I and a type III protease of

CC L.lactis Wg2 and SK11.

CC The product has modified properties, e.g. thermostability,

CC alkaline/acid pH stability, oxidative stability, autolysis etc.,

CC compared to the parent protease(s). The proteases can be used for

CC preparing products (butter cheese, human and animal foodstuffs)

CC prepared with the aid of lactic acid bacteria.

CC See also Q10411-17 and Q10870-71.

XX Sequence 1962 AA;

XX Query Match 39.6%; Score 40; DB 12; Length 1962;

XX Best Local Similarity 35.0%; Pred. No. 4.6e+02;

XX Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 IMODDPQNSIEATSVKSKY 20

: || | | : ||

Db 1414 vytndpnfqtgtatdnaqy 1433

PT foodstuffs and flavourings

XX Disclosure; Fig 1(1-7)+5(a); 29 pp; English.

XX The mutant protease N166D having new cleavage specificities is obtained by carrying out single amino acid substitutions.

CC This mutant may then be used to prepare hybrid proteases, the fusion being between a type I and a type III protease of

CC L.lactis Wg2 and SK11.

CC The product has modified properties, e.g. thermostability, alkaline/acid pH stability, oxidative stability, autolysis etc., compared to the parent protease(s). The proteases can be used for preparing products (butter cheese, human and animal foodstuffs) prepared with the aid of lactic acid bacteria.

CC See also Q10411-17 and Q10870-71.

XX Sequence 1962 AA;

Query Match 39.6%; Score 40; DB 12; Length 1962;

Best Local Similarity 35.0%; Pred. No. 4.6e+02;

Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 IMQDDPQNSIEATSVDSQY 20

: :||| | | :||

Db 1414 vytndpnfgitgtatdnagy 1433

RESULT 9

RI0557
ID RI0557 standard; Protein; 1962 AA.

XX AC RI0557;

XX DT 15-APR-1991 (first entry)

XX DE Mutant protease (A137G/K138D).

XX KW Mutant protease gene; fermentation; foodstuff; flavouring; lactic acid bacteria.

XX OS Lactococcus lactis SK11.

XX FH Key Location/Qualifiers

XX FT Peptide 1..187

XX FT /label= signal_peptide

XX FT 188..1962

XX FT /label= mature_protein

XX PN EP411715-A.

XX PD 06-FEB-1991.

XX PF 02-AUG-1990; 90EP-0202113.

XX PR 04-AUG-1989; 89NL-0002010.

XX PA (NEZU-) NED INST ZUIVELONDE.

XX PI Vos PAJ, Siezen RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ;

XX DR WPI; 1991-038622/06.

XX DR N-PSDB; Q10411.

XX PT Mutant protease gene(s) and protease(s) - derived from type I and III protease genes from lactococcal strains, used in fermentation foodstuffs and flavourings

XX Disclosure; Fig 1(1-7)+5(b); 29 pp; English.

XX The mutant protease A137G/K138D having new cleavage specificities is obtained by replacing two amino acids.

CC This mutant may then be used to prepare hybrid proteases,

CC the fusion being between a type I and a type III protease of L.lactis Wg2 and SK11.

CC The product has modified properties, e.g. thermostability, alkaline/acid pH stability, oxidative stability, autolysis etc., compared to the parent protease(s). The proteases can be used for preparing products (butter cheese, human and animal foodstuffs) prepared with the aid of lactic acid bacteria.

CC See also Q10411-17 and Q10870-71.

XX Sequence 1962 AA;

XX Query Match 39.6%; Score 40; DB 12; Length 1962;

XX Best Local Similarity 35.0%; Pred. No. 4.6e+02;

XX Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

XX Qy 1 IMQDDPQNSIEATSVDSQY 20

: :||| | | :||

XX Db 1414 vytndpnfgitgtatdnagy 1433

XX RESULT 10

XX RI0558

XX ID RI0558 standard; Protein; 1962 AA.

XX AC RI0558;

XX DT 15-APR-1991 (first entry)

XX DE Mutant protease (A137G/K138P/T139P).

XX KW Mutant protease gene; fermentation; foodstuff; flavouring; lactic acid bacteria.

XX OS Lactococcus lactis SK11.

XX FH Key Location/Qualifiers

XX FT Peptide 1..187

XX FT /label= signal_peptide

XX FT 188..1962

XX FT /label= mature_protein

XX PN EP411715-A.

XX PD 06-FEB-1991.

XX PF 02-AUG-1990; 90EP-0202113.

XX PR 04-AUG-1989; 89NL-0002010.

XX PA (NEZU-) NED INST ZUIVELONDE.

XX PI Vos PAJ, Siezen RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ;

XX DR WPI; 1991-038622/06.

XX DR N-PSDB; Q10412.

XX PT Mutant protease gene(s) and protease(s) - derived from type I and III protease genes from lactococcal strains, used in fermentation foodstuffs and flavourings

XX Disclosure; Fig 1(1-7)+5(b); 29 pp; English.

XX The mutant protease having new cleavage specificities is obtained by replacing three amino acids.

CC This mutant may then be used to prepare hybrid proteases, the fusion being between a type I and a type III protease of

CC L.lactis Wg2 and SK11.

CC The product has modified properties, e.g. thermostability, alkaline/acid pH stability, oxidative stability, autolysis etc., compared to the parent protease(s). The proteases can be used for

CC preparing products (butter cheese, human and animal foodstuffs) prepared with the aid of lactic acid bacteria.

XX PF 02-AUG-1990; 90EP-0202113.
 XX PR 04-AUG-1989; 89NL-0002010.
 XX PA (NEZU-) NED INST ZUIVELONDE.
 XX PI Vos PAJ, Siezen RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ;
 XX DR WPI; 1991-038622/06.
 XX DR N-PSDB; Q10416.
 XX PT Mutant protease gene(s) and protease(s) - derived from type I and
 PT III protease genes from lactococcal strains, used in fermentation
 PT foodstuffs and flavourings
 XX PS Disclosure: Fig 1(1-7)+5(b); 29 pp; English.
 XX CC The mutant protease having new cleavage specificities is
 CC obtained by deleting three amino acids.
 CC CC This mutant may then be used to prepare hybrid proteases,
 CC the fusion being between a type I and a type III protease of
 CC L.lactis Wg2 and SK11.
 CC CC The product has modified properties, e.g. thermostability,
 CC alkaline/acid pH stability, oxidative stability, autoprolysis etc.,
 CC compared to the parent protease(s). The proteases can be used for
 CC preparing products (butter cheese, human and animal foodstuffs)
 CC prepared with the aid of lactic acid bacteria.
 CC See also Q10411-17 and Q10870-71.
 XX SQ Sequence 1959 AA;

Query Match 39.6%; Score 40; DB 12; Length 1959;
 Best Local Similarity 35.0%; Pred. No. 4.6e+02;
 Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 1 IMODDPONSIEATSVKRSQY 20
 : :||| | | :|||
 Db 1411 vytndpnfqtgtatdnaqy 1430

RESULT 7
 RI0560
 ID RI0560 standard; Protein; 1962 AA.
 XX AC RI0560;
 XX DT 15-APR-1991 (first entry)
 XX DE Mutant protease (K138N).
 XX KW Mutant protease gene; fermentation; foodstuff; flavouring;
 KW lactic acid bacteria.
 XX OS Lactococcus lactis SK11.
 XX FH Key Location/Qualifiers
 FT Peptide 1..187
 FT Protein /label= signal_peptide
 FT 188..1962
 FT /label= mature_protein
 XX PN EP411715-A.
 XX PD 06-FEB-1991.
 XX PF 02-AUG-1990; 90EP-0202113.
 XX PR 04-AUG-1989; 89NL-0002010.
 XX PA (NEZU-) NED INST ZUIVELONDE.
 XX PT

PI Vos PAJ, Siezen RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ;
 XX DR WPI; 1991-038622/06.
 XX DR N-PSDB; Q10414.
 XX PT Mutant protease gene(s) and protease(s) - derived from type I and
 PT III protease genes from lactococcal strains, used in fermentation
 PT foodstuffs and flavourings
 XX PS Disclosure: Fig 1(1-7)+5(a); 29 pp; English.
 XX CC The mutant protease K138N having new cleavage specificities is
 CC obtained by carrying out single amino acid substitutions.
 CC CC This mutant may then be used to prepare hybrid proteases,
 CC the fusion being between a type I and a type III protease of
 CC L.lactis Wg2 and SK11.
 CC CC The product has modified properties, e.g. thermostability,
 CC alkaline/acid pH stability, oxidative stability, autoprolysis etc.,
 CC compared to the parent protease(s). The proteases can be used for
 CC preparing products (butter cheese, human and animal foodstuffs)
 CC prepared with the aid of lactic acid bacteria.
 CC See also Q10411-17 and Q10870-71.
 XX SQ Sequence 1962 AA;

Query Match 39.6%; Score 40; DB 12; Length 1962;
 Best Local Similarity 35.0%; Pred. No. 4.6e+02;
 Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 1 IMODDPONSIEATSVKRSQY 20
 : :||| | | :|||
 Db 1414 vytndpnfqtgtatdnaqy 1433

RESULT 8
 RI0561
 ID RI0561 standard; Protein; 1962 AA.
 XX AC RI0561;
 XX DT 15-APR-1991 (first entry)
 XX DE Mutant protease (N166D).
 XX KW Mutant protease gene; fermentation; foodstuff; flavouring;
 KW lactic acid bacteria.
 XX OS Lactococcus lactis SK11.
 XX FH Key Location/Qualifiers
 FT Peptide 1..187
 FT Protein /label= signal_peptide
 FT 188..1962
 FT /label= mature_protein
 XX PN EP411715-A.
 XX PD 06-FEB-1991.
 XX PF 02-AUG-1990; 90EP-0202113.
 XX PR 04-AUG-1989; 89NL-0002010.
 XX PA (NEZU-) NED INST ZUIVELONDE.
 XX PI Vos PAJ, Siezen RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ;
 XX DR WPI; 1991-038622/06.
 XX DR N-PSDB; Q10415.
 XX PT Mutant protease gene(s) and protease(s) - derived from type I and
 PT III protease genes from lactococcal strains, used in fermentation

```

XX
SQ Sequence 724 AA;

Query Match 40.6%; Score 41; DB 13; Length 724;
Best Local Similarity 37.5%; Pred. No. 96;
Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 5 DPONSIEATSVDKSQY 20
Db 114 dpdptvkttkisksqf 129
||:::|:||||:

RESULT 4
R86558
ID ID R86558 standard; Protein; 256 AA.
XX
XX
AC R86558;
XX
XX 19-FEB-1996 (first entry)
XX
XX Mouse OBF-1.
XX
XX
KW Oct binding factor 1; octamer site-mediated gene transcription;
KW B-lymphocyte specific activator; OBF-1; POU protein; oct-1; oct-2.
XX
XX Mus sp.
XX
XX WO9532284-A1.
XX
XX 30-NOV-1995.
XX
XX 15-MAY-1995; 95WO-EP01834.
XX
XX 24-MAY-1994; 94EP-0810299.
XX
XX (CIBA ) CIBA GEIGY AG.
XX
XX Matthias P, Strubin M;
XX
XX WPI; 1996-020579/02.
XX
XX N-PSDB; T06608.
XX
XX DNA encoding B-lymphocyte specific activator of octamer
XX site-mediated gene transcription - used to identify proteins
XX interacting with POU proteins Oct-1 and Oct-2
XX
XX Claim 15; Page 47-48; 55pp; English.
XX
XX The mouse homologue (see T06608) of human OBF-1 cDNA was isolated
XX from a cDNA library prepd. from the mouse B-cell line S194 by
XX homology hybridization. The cDNA is used to express mouse OBF-1
XX (R86558) in heterologous cells for use in identifying potential
XX drugs that modulate OBF-1 activity.
XX
XX Sequence 256 AA;

Query Match 39.6%; Score 40; DB 17; Length 256;
Best Local Similarity 42.9%; Pred. No. 41;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 4 DDPONSIEATSVDK 17
Db 221 ddpraissitdk 234
||||:|:||||

RESULT 5
P94145
ID ID P94145 standard; protein; 1560 AA.
XX
XX AC P94145;
XX

```

PT hyper-calcaemia, hypo- or hyper-phosphatemia, kidney stones, etc
 XX
 PS Claim 9; Fig 1A-E; 62pp; English.
 CC
 CC A novel 7-transmembrane receptor (W12695) has been identified as a
 CC human G-protein parathyroid hormone (PTH) receptor, designated
 CC HTD674. It shows 48.2% homology to the human PTH receptor. Its
 CC amino acid sequence was deduced from a cDNA clone (759619) isolated
 CC from a human T cell lymphoma tissue cDNA library. Recombinant
 CC HTD674 can be produced in transformed host cells and used to
 CC screen for (ant)agonist cpds. Agonists may be used to prevent or
 CC treat e.g. hypocalcaemia, hyperphosphataemia, hypoparathyroidism
 CC and chronic tetany by stimulating an increase in serum calcium
 CC levels. Antagonists can be used to inhibit the receptor e.g. for
 CC the treatment of osteoporosis, hypercalcaemia, hypoparathyroidism,
 CC hypophosphataemia, kidney stone, nephrolisis.
 XX
 SQ Sequence 541 AA;

Query Match 100.0%; Score 101; DB 18; Length 541;
 Best Local Similarity 100.0%; Pred. No. 6e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IMQDDPQNSIEATSVKXSY 20
 |||||
 Db 214 imqddpqnsieatsvksqy 233

RESULT 2
 W38757
 ID W38757 standard; Protein; 1876 AA.
 AC W38757;
 XX
 DT 22-JUN-1998 (first entry)
 XX
 XX Phosphatidyl inositol 3-kinase cdk.
 DE
 XX Phosphatidyl inositol 3-kinase; signal transduction; cell cycle;
 KW antagonist; inflammatory joint disease; cell proliferation; cancer;
 KW psoriasis; restenosis; atherosclerosis; therapy; diagnosis; cdk.
 OS
 XX Drosophila melanogaster.
 XX
 FH Key Location/Qualifiers
 FT Peptide 420..434
 FT /note= "Claim 10"
 FT Domain 1740..1876
 FT /note= "C2 domain"
 FT Domain 1324..1594
 FT /note= "catalytic domain"
 XX
 PN W09731650-A1.
 XX
 PD 04-SEP-1997.
 XX
 XX 12-FEB-1997; 97WO-US02193.
 XX
 PR 29-FEB-1996; 96US-0609049.
 XX
 XX (REGC) UNIV CALIFORNIA.
 PA
 XX Chen Y, Molz L, Williams LT;
 XX
 XX WPI; 1997-44842/41.
 DR N-PSDB; T80200.
 XX
 XX New isolated phosphatidyl inositol-3 kinase polypeptide - used to
 PT develop products for diagnosis and therapy, particularly for
 PT proliferative disorders, e.g. inflammatory joint diseases, or cancer
 XX
 XX Claim 5; Fig 10; 77pp; English.

XX This protein sequence comprises cpk, a Drosophila polypeptide that
 CC belongs to a novel class of phosphatidyl inositol 3-kinases that
 CC contain a C2 domain, are capable of phosphorylating a D3 hydroxyl
 CC of an inositol ring in phosphatidyl inositol and phosphatidyl
 CC inositol 4-phosphate, but not in phosphatidyl inositol
 CC 4,5-bisphosphate, and which are involved in cell signalling
 CC cascades that control e.g. cell cycle progression and intracellular
 CC protein sorting. The amino acid sequence was deduced from an
 CC isolated cDNA sequence (see T80200). It shows 34% identity and
 CC 48% similarity to mouse cpk-m (see W38756). Novel phosphatidyl
 CC inositol 3-kinases can be used to screen for agonists/antagonists
 CC of activity and in a claimed method of treating a disorder caused
 CC by dysregulation of a growth factor activation signalling cascade.
 CC Antagonists may reduce Ras activation allowing treatment of
 CC proliferative disorders such as atherosclerosis, inflammatory joint
 CC disease, psoriasis, restenosis following angioplasty, and cancer.
 XX
 SQ Sequence 1876 AA;

Query Match 45.5%; Score 46; DB 18; Length 1876;
 Best Local Similarity 56.2%; Pred. No. 43;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 MQDDPQNSIEATSVK 17
 : ||| ||| | ||:
 Db 1187 lpdqphnsigaamvdq 1202

RESULT 3
 R29002
 ID R29002 standard; Protein; 724 AA.
 XX
 AC R29002;
 XX
 DT 04-JAN-1993 (first entry)
 XX
 DE Protein derived from ORF3 of pSRQ220.
 DE
 XX Listeria; Pedococcus; precursor; ORF; food.
 KW
 XX Pedicoccus acidilactici.
 OS
 XX EP493779-A.
 PN
 XX 08-JUL-1992.
 PD
 XX 23-DEC-1991; 91EP-0122124.
 PF
 XX 31-DEC-1990; 90US-0635965.
 PR
 XX (QUES-) QUEST INT FLAVORS & FOODS INGREDIENTS.
 PA (QUES-) QUEST INT BV.
 PA
 XX Henderson JT, Ledeboer AM, Marugg JD, Vanderbergh PA;
 PI
 XX WPI; 1992-227391/28.
 DR N-PSDB; Q25810.
 DR
 XX Cloned gene from Pedicoccus acidilactici - encodes bacteriocin
 PT used to inhibit Listeria in foods
 PT
 XX Disclosure; Fig 4; 44pp; English.
 PS
 XX The protein sequence was deduced from the third ORF encoded by the
 CC 5.6 kbp EcoRI-SalI insert from pSRQ220. This vector contains three
 CC ORFs, the first encoding a bacteriocin precursor, and the others
 CC needed for correct protein assembly to obtain functionally active
 CC bacteriocin. The bacteriocin from Pedicoccus acidilactici NRRL-B-
 CC 18050 is used in foods to inhibit bacterial spoilage, e.g. to inhibit
 CC Listeria monocytogenes.
 CC See also R25810 and R29001.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:51:41 ; Search time 138.73 Seconds
(without alignments)
4.930 Million cell updates/sec

Title: US-09-236-468A-2_COPY_214_233
Perfect score: 101
Sequence: 1 IMODDPQNSIETATSVKSOY 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_36.*

1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	541	18	G-protein parathy
2	46	45.5	1876	18	Phosphatidyl inosi
3	41	40.6	724	13	Protein derived fr
4	40	39.6	256	17	Mouse OBF-1. Mus
5	40	39.6	1560	10	S. cremoris protei
6	40	39.6	1959	12	Mutant protease (d
7	40	39.6	1962	12	Mutant protease (K
8	40	39.6	1962	12	Mutant protease (N
9	40	39.6	1962	12	Mutant protease (A
10	40	39.6	1962	12	Mutant protease (A
11	40	39.6	1962	12	Mutant protease (A
12	40	39.6	1962	12	Mutant protease (K

13	40	39.6	1968	12	R10941	Mutant protease (d
14	40	39.6	1974	12	R10940	Mutant protease (d
15	39	38.6	233	12	R14361	Immunogenic serine
16	39	38.6	471	16	R67692	C. albicans caauri
17	39	38.6	471	17	R88133	Aureobasidin sensi
18	39	38.6	472	16	R67696	C. albicans caauri
19	38	37.6	75	19	W42018	Clone BL341_4 prot
20	38	37.6	75	20	Y08625	Human secreted pro
21	38	37.6	75	21	Y67316	Human secreted pro
22	38	37.6	80	19	Y86066	S. pneumoniae deri
23	38	37.6	82	18	W28194	Amino acid sequenc
24	38	37.6	136	19	W44691	Staphylokinase mut
25	38	37.6	136	20	Y15037	Staphylokinase var
26	38	37.6	136	20	Y15014	Staphylokinase var
27	38	37.6	373	21	Y75199	Neisseria gonorrhoe
28	38	37.6	456	14	R39902	B. canis 21B4/thop
29	38	37.6	558	21	Y75200	Neisseria meningit
30	38	37.6	558	21	Y75201	Neisseria meningit
31	38	37.6	661	11	R07504	Merzoite apical-en
32	38	37.6	662	18	W24576	Merzoite apical-e
33	37.5	37.1	512	19	W31554	Collagen binding p
34	37.5	37.1	1185	13	R22675	Collagen binding p
35	37	36.6	123	14	R39152	Staphylokinase SAK
36	37	36.6	126	13	R25468	Thrombolytic pepti
37	37	36.6	127	14	R39151	Staphylokinase SAK
38	37	36.6	136	17	W03077	Staphylokinase der
39	37	36.6	136	17	W03078	Staphylokinase der
40	37	36.6	136	17	W03082	Staphylokinase der
41	37	36.6	136	17	W03083	Staphylokinase der
42	37	36.6	136	17	W03075	Staphylokinase der
43	37	36.6	136	17	W03076	Staphylokinase der
44	37	36.6	136	17	W03098	Staphylokinase der
45	37	36.6	136	21	Y50871	Staphylococcus aur

ALIGNMENTS

RESULT 1

W12895
ID W12695 standard; Protein; 541 AA.

XX W12695;

XX 31-MAY-1997 (first entry)

XX G-protein parathyroid hormone receptor HLTGD74.

DE G-protein parathyroid hormone receptor; HLTGD74; parathormone; PTH;
XX calcium; signal transduction; agonist; antagonist; hypocalcaemia;
KW hyperphosphataemia; hypoparathyroidism; chronic tetany;
KW osteoporosis; hypercalcaemia; hypoparathyroidism; hypophosphataemia;
KW kidney stone; nephrolithiasis; therapy; diagnosis.

XX Homo sapiens.

XX WO9639433-A1.

XX 12-DEC-1996.

XX 05-JUN-1995; 95WO-US07085.

XX 05-JUN-1995; 95WO-US07085.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Li Y, Rosen CA, Ruben SM, Soppet DR;

XX WPI: 1997-043068/04.

XX N-PSDB: t59619.

XX Human G-protein parathyroid hormone receptor, HLTGD74 - used to
PT identify (ant)agonists, used in the treatment of hypo- or

AC	09TUJ1;
AD	01-MAY-2000 (TrEMBLrel. 13, Created)
AE	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
AF	01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
AG	GROWTH HORMONE-RELEASING HORMONE RECEPTOR SHORT FORM.
AH	Bos taurus (Bovine).
AI	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AJ	OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
AK	Bovidae; Bovinae; Bos.
AL	[1]
AM	SEQUENCE FROM N.A.
AN	RP Takata M., Tarumi O., Watanabe S., Sekikawa K.;
AO	RA "Molecular cloning of bovine growth hormone-releasing hormone receptor
AP	RT CDNA."
AQ	RT Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
AR	RL EMBL; AS022596; BAA84959.1; -.
AS	DR INTERPRO: IPR000832; -.
AT	DR INTERPRO: IPR001771; -.
AV	DR INTERPRO: IPR001879; -.
AW	DR PFAM; PF00002; 7tm.2; 1.
AX	DR PRINTS; PR00249; GPCRSECRETIN.
AY	DR PRINTS; PR01154; VIPIRECEPTOR.
AZ	DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
BA	DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
BB	DR Receptor.
BC	DR KW
BD	DR
BE	DR
BF	DR
BG	DR
BH	DR
BI	DR
BJ	DR
BK	DR
BL	DR
BM	DR
BN	DR
BO	DR
BP	DR
BQ	DR
BR	DR
BS	DR
BT	DR
BU	DR
BV	DR
BW	DR
BX	DR
BY	DR
BZ	DR
CA	DR
CB	DR
CC	DR
CD	DR
CE	DR
CF	DR
CG	DR
CH	DR
CI	DR
CJ	DR
CK	DR
CL	DR
CM	DR
CN	DR
CO	DR
CP	DR
CQ	DR
CR	DR
CS	DR
CT	DR
CU	DR
CV	DR
CW	DR
CX	DR
CY	DR
CA	DR
CB	DR
CC	DR
CD	DR
CE	DR
CF	DR
CG	DR
CH	DR
CI	DR
CJ	DR
CK	DR
CL	DR
CM	DR
CN	DR
CO	DR
CP	DR
CQ	DR
CR	DR
CS	DR
CT	DR
CU	DR
CV	DR
CW	DR
CX	DR
CY	DR
CA	DR
CB	DR
CC	DR
CD	DR
CE	DR
CF	DR
CG	DR
CH	DR
CI	DR
CJ	DR
CK	DR
CL	DR
CM	DR
CN	DR
CO	DR
CP	DR
CQ	DR
CR	DR
CS	DR
CT	DR
CU	DR
CV	DR
CW	DR
CX	DR
CY	DR
CA	DR
CB	DR
CC	DR
CD	DR
CE	DR
CF	DR
CG	DR
CH	DR
CI	DR
CJ	DR
CK	DR
CL	DR
CM	DR
CN	DR
CO	DR
CP	DR
CQ	DR
CR	DR
CS	DR
CT	DR
CU	DR
CV	DR
CW	DR
CX	DR
CY	DR
CA	DR
CB	DR
CC	DR
CD	DR
CE	DR
CF	DR
CG	DR
CH	DR
CI	DR
CJ	DR
CK	DR
CL	DR
CM	DR
CN	DR
CO	DR
CP	DR
CQ	DR
CR	DR
CS	DR
CT	DR
CU	DR
CV	DR
CW	DR
CX	DR
CY	DR
CA	DR
CB	DR
CC	DR
CD	DR
CE	DR
CF	DR
CG	DR
CH	DR
CI	DR
CJ	DR
CK	DR
CL	DR
CM	DR
CN	DR
CO	DR
CP	DR
CQ	DR
CR	DR
CS	DR
CT	DR
CU	DR
CV	DR
CW	DR
CX	DR
CY	DR
CA	DR
CB	DR
CC	DR
CD	DR
CE	DR
CF	DR
CG	DR
CH	DR
CI	DR
CJ	DR
CK	DR
CL	DR
CM	DR
CN	DR
CO	DR
CP	DR
CQ	DR
CR	DR
CS	DR
CT	DR
CU	

Query Match	23.6%	Score	685.5;	DB	13;	Length	465;
Best Local Similarity	32.1%	Pred.	No. 1.9e-51;				
Matches	157;	Conservative	97;	Mismatches	168;	Indels	67;
Gaps	14;						
Qy	13	WLMGSCLLARAOLDSTITIEQIVLVLKAKVQ-CFLNITAOLOEGEGNCFPEWQGLI	71				
Db	8	FLLLGFGMASQVASHPHYCIILKEEACL--EKIQRYEIEMWNTQSG---CPGMWONIT	61				
Qy	72	CWPRGTGKGISAVPCPPY-----IYDFNKHGVAFRHCNPNGT	108				
Db	62	CWMAPEVGKVVSRCPALFSMIGSEDEMDFVDRSLGWSPENIEQSQEGTKIRNCTENG-	120				
Qy	109	WDFMHSLNKWTWANSYDCLRFLOPDISIGKEFCERLYVMTVGVYSIFSGSLAVAILIGY	168				
Db	121	W-----SEPPPHYSEACDFDINTGPDQDQYIYLSVKALYTVGVYSTSLVALTAMVILCR	174				
Qy	169	FRRLHCTRNTHMHLFVSFMLRATSFVKQVRVVAHIGVKLESLIMQDDPQNSIEATSV	228				
Db	175	FRKLHCTRNTHNHLFVSFILRAISVFIKDEVLYAE-----QDN--NICHLSV	221				
Qy	229	DKSOYICGCTAVVMYFIYELATNYWIIIVEGLYLHNLIFVAFSTKYLMGFIILGWGPPA	288				
Db	222	E-----CKVMVVFHYCVMSNYEFLFIEGLYLFLELVTFVFPERRFYFWYTIIGWGTP	275				
Qy	289	AFVAAMAVARATLADARCWELSAG-DKWKIYQAPILAAIGLNFILFNTVRLVLTAKIKWET	347				
Db	276	ICVTIWAVALRHLFDNLGCDWDTNNNTGLWWYIKGPVIGSIMINFVLFGVIGIILVQKL-QS	334				
Qy	348	NAVGHDRKQYRKUAKSTLVLLVFGVHHYIVFVCLPHSFTGLGWEIFRMHCELPFNSFOGF	407				
Db	335	PDJGGNESSIYLRARSTLLPIUFGHYTFAFSPEN---VSKRERLVFLFELGSGFOGF	391				

Query Match	22.9%	Score	665;	DB	11;	Length	550;
Best Local Similarity	34.5%	Pred.	No. 1.4e-49;				
Matches	152;	Conservative	76;	Mismatches	156;	Indels	56;
Gaps	15;						
Qy	32	ITIEEQIVLVIAKVOCELNITAOIQEGCN-----CFPEWDCGLICWPRTCVCKISAVP	85				
Db	59	VSIAQVGTSLUKETQKWANTYKCELBHNLRSIGFCNGTFDRVCMPHSYGNVS--VP	117				
Qy	86	CPPIY--YDFNHKGVAFRHCNPGNTWDFMHSNLKNTWANSDCLR--FLOPDIISGKQEF	140				
Db	118	CPSYLPWNAESPGRYRHCLAQCTWQTRENTDIWDQSECSENHSFRQ---NVDHYAL	174				
Qy	141	CERLYVMYTVGYSTISFGSLAVAILIIGYFRRLHCTRNYIHMHILVSPMLRATSTFVKDRV	200				
Db	175	LYTIQLMYTVGYSYSLISFLTALTFLFLRKLHCTRNYIHMLNFASFETKLVALVLVMDV	234				
Qy	201	VHAHIGVKELSLIMQDDPONSIEATSVDKSOY-----IGCKIAVMVFIFYLATNYY	252				
Db	235	SH-----NSYKRPDDSEGWSYLSSETSVCSCRVOQLHYFVGNHL	276				
Qy	253	WILVEGYLHNLIFVAFSDPKYLV-GLILGWGPPAAFAVAAMARATLADARCWLSA	311				
Db	277	WLLVEGYLVHTLLEPTVPERR-LWPKYLVVGVGAFFPMLFVLPWGFAHLENTRCWATN-	334				
Qy	312	GDIK--WIYQAPILAAIGLNLFLNTVRVLATKIWETNAVGHDTROYKRLAKSTLVLV	369				
Db	335	GNLKIWIIRGPMLLCVTVNFEFFLKLKLKLSIKLKAHQMCFDYK--YR-LAKSTLLI	391				
Qy	370	LVFGVHYIVFVCLP-HSGFTGLGWIRMHCELFNFSFGGFVSYLIYCNCGEQVAEVKMW	428				
Db	392	PLLGVHEVLEFTFFDDOVGSEKRIPIOLTLSSVHGICFLVALOYGVANGVAKRLKSW	451				

```
Db 65 KWNLDNLTCTWPTTPGWQVVLDCPLIFQLFSPHNGVNSRNCTEEG-----WSQ- 112
QY 123 SDCRLFLQP---DISIG-----KQFCERLVMYMTVGYSISFGSLAVAILIG 167
Db 113 -----LEPGPYHACGLNDRASSMDEQOQTEFYDAVKTYGTIGYISLSLASLLVAMAIL 166
QY 168 YPRRLHCTRNYTHMLFVSFMLRATSFVKDRVVAHIGVKELESIMODDPQNSIEATS 227
Db 167 LPRKRLHCTRNYTHMLFVSFMLRATAVFKDM-----ALFNNGETDHCSEAS- 213
QY 228 VDKSQYICKIAVVMFYFLATNYWILVEGLYLNHLIFVAFESDTKYLWGFIILGWGFP 287
Db 214 -----VSKAAVFPQCVCWANEFWLVEGLYLNHLIAVFFSSEKRYFWGYIILGWGVP 267
QY 288 AAFVAAVARATLADARCWEL-LSAGDIKIWIYQAPILAAIGNFIILFNTVRLVATKIWE 346
Db 268 SVFIMWTIVRIHPEDFCWDFTIINSSLMWIKIGPILISILVNFILFICIRILVQKLRP 327
QY 347 TNAVGHDRKQYRKLAKTSLVLVLFVGVHYIVFVCLPHSFTGLGWEIRMHCELFNFSQ 406
Db 328 PD-IGKNDSSPYSRUAKSTLLIPLFGVHYVMAFPDFNFKRA--QVKMVFELVVGSPQ 383
QY 407 FVSIYCYCNGEVOAEVKKWWSRWNL-VDW--KRTPPCGSRR--CG---SVLTVT 456
Db 384 FVAILCYFLNGEVOAEVKKWWSRWNLHGVLGWSSKQHPWGGNGVSCSTQVSNLTRVS 443
QY 457 ----HSTSQSQVA 466
Db 444 PSARSSSFQAEVS 457
Db 444 PSARSSSFQAEVS 457
RESULT 8
ID 073769 PRELIMINARY; PRT; 465 AA.
AC 073769;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE PITUITARY ADENYLATE CYCLASE ACTIVATING POLYPEPTIDE TYPE 1 RECEPTOR
DE PRECURSOR
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprinidae; Cyprinidae; Carassius.
RN [1]
RP SEQUENCE FROM N.A.
RA Wong A.O.L., Leung M.Y., Shea W.L.C., Chang J.P., Chow B.K.C.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF048820; AAC15699.1; -.
DR INTERPRO; IPR000832; -.
DR INTRERPRO; IPR001879; -.
DR PFAM; PF00002; 7tm-2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
SQ SEQUENCE 465 AA; 53459 MW; 9779A95EDBFD1DC1 CRC64;
Query Match 24.58; Score 713; DB 13; Length 465;
Best Local Similarity 35.6%; Pred. No. 7.7e-54;
Matches 156; Conservative 90; Mismatches 138; Indels 54; Gaps 16;
QY 58 EGEKCPPEWDLICWPRGTGKISAVCPDPIYDF----NHKGVAFRHCNPNCTWDFMH 113
Db 65 DGKPECPWEWDLICWPRGTGKISAVCPDPIYDF----NHKGVAFRHCNPNCTWDFMH 118
QY 114 SLNKTWANSYD-CL---RFLQDISIGKQECERLYVMYTVGVSGISFGSLAVAILIIGY 169
Db 119 --SESYPHYVDACMIGENTTKPDN-----YYASVKALYTVGYSTSLVSLTANVILCRF 170
QY 170 RRLHCTRNYTHMLFVSFMLRATSFVKDRVVAHIGVKELESIMODDPQNSIEATSVD 229
Db 171 RKLHCTRNYTHMLFVSFMLRATSFVKDRVVAHIGVKELESIMODDPQNSIEATSVD 229
Db 171 RKLHCTRNYTHMLFVSFMLRATSFVKDRVVAHIGVKELESIMODDPQNSIEATSVD 229
```

```
QY 230 KSOYIGCKIAVVMFYFLATNYWILVEGLYLNHLIFVAFESDTKYLWGFIILGWGFPAA 289
Db 216 ---VGCKAVVMFYFLATNYWILVEGLYLNHLIFVAFESDTKYLWGFIILGWGFPAA 271
QY 290 FVAWAVARATLADARCWELSGD- IKWIYQAPILAAIGNFIILFNTVRLVATKIWE 348
Db 272 CVTIWAVLRLHFDGSGCWMNDNTALWVWIKGPPVVASIMINFLVFIGIILVQKL-QSP 330
QY 349 AVGHDRKQYRKLAKTSLVLVLFVGVHYIVFVCLPHSFTGLGWEIRMHCELFNFSQGF 408
Db 331 DIGGNESSIYLRARSTLLIPLFGVHYVMAFPDFNFKRA--QVKMVFELVVGSPQ 387
QY 409 VSIYCYCNGEVOAEVKKWWSRWNL-VDW--KRTPPCGSRR--CG---SVLTVT 461
Db 388 VAVLYCYFLNGEVOAEVKKWWSRWNLHGVLGWSSKQHPWGGNGVSCSTQVSNLTRVS 447
QY 462 QSQVAAAHAWCLSLAKLP 479
Db 448 QIRMSSPLAETVNL-NLP 464
RESULT 9
ID 073768 PRELIMINARY; PRT; 438 AA.
AC 073768;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE GROWTH-HORMONE RELEASING HORMONE-LIKE PEPTIDE RECEPTOR.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprinidae; Cyprinidae; Carassius.
RN [1]
RP SEQUENCE FROM N.A.
RA Chan K.W., Yu K.L., Rivier J., Chow B.K.C.;
RL Neuroendocrinology 0:0-0(1998).
DR EMBL; AF048819; AAC15698.1; -.
DR INTERPRO; IPR000832; -.
DR INTRERPRO; IPR001879; -.
DR PFAM; PF00002; 7tm-2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; UNKNOWN_1.
SQ SEQUENCE 438 AA; 50305 MW; C3516A52D692C339 CRC64;
Query Match 24.2%; Score 702.5; DB 13; Length 438;
Best Local Similarity 35.5%; Pred. No. 5.9e-53;
Matches 148; Conservative 81; Mismatches 147; Indels 41; Gaps 11;
QY 57 QEGKCPPEWDLICWPRGTGKISAVCPDPIYDFNHKGVAFRHCNPNCTWDFMHSL 115
Db 55 QSOQTGCTWDGIRCWQAKTAKTGLINVSQSDVQFHSINTQGFYRNCSSNG-W----- 107
QY 116 NKTWANSYDCLRELQPDISIGKQECERLYVMYTVGVSGISFGSLAVAILIIGYFRLHCT 175
Db 108 SDPIPTEEACTFEDDSSEGTSTYLSLTKOLYTAGYATSLISLITAVIPTCFKPHCT 167
QY 176 RNYTHMLFVSFMLRATSFVKDRVVAHIGVKELESIMODDPQNSIEATSVKQSYIG 235
Db 168 RNYTHMLFVSFMLRATSFVKDRVVAHIGVKELESIMODDPQNSIEATSVKQSYIG 235
QY 236 CKIAVVMFYFLATNYWILVEGLYLNHLIFVAFESDTKYLWGFIILGWGFPAAFAAWA 295
Db 209 CKTAVTFQFCILTFEWMFLVEGLYLTILTLTFVSRQKRYFMWYILIGWGVSVVLIWV 268
QY 296 VARATLADARCWELSGD-DIKWIYQAPILAAIGNFIILFNTVRLVATKIWE 354
Db 269 LTRQFYDNRCGWDNTDNNIIMWIKGPPVVASIMINFLVFIGIILVQKLKSPGVGND 328
QY 355 RKQYRKLAKTSLVLVLFVGVHYIVFVCLPHSFTGLGWEIRMHCELFNFSQGFVSIYC 414
```

```

QY 250 NYWILVEGLYLNHLNIFVAFESDFTKYUWGFILIGWGPAAFAVAAMAVARATLADARCWEL 309
Db 1 NYWILVEGLYLNHLNIFVAFESDFTKYUWGFILIGWGPAAFAVAAMAVARATLADARCWEL 60
QY 310 SAGDIKIYQAPILAAIGLNFILFNTVRVLTAKIWTETNAVGHDTKQYRKLAQSTLVLY 369
Db 61 SAGD-RWYQAPILAAIGLNFILFNTVRVLTAKIWTETNAVGHDTKQYRKLAQSTLVLY 119
QY 370 LVFGVHYIVFVCLPHSFTGLGWEIRHCELFNFSGFFVSIYCYCNGE 419
Db 120 LVFGVHYIVFVCPHSPHSGLWELHCELFNFSGFFVSIYCYCNGE 169

RESULT 6
Q9YHC6
ID Q9YHC6 PRELIMINARY; PRT; 444 AA.
AC Q9YHC6;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE VASOACTIVE INTESTINAL PEPTIDE/PITUITARY ADENYLATE CYCLASE ACTIVATING
DE POLYPEPTIDE RECEPTOR.
OS Rana ridibunda (Laughing frog) (Marsh frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PITUITARY;
RA Alexandre D., Anouar Y.;
RT "A cloned frog VIP/PACAP receptor exhibits pharmacological and tissue
RT distribution characteristics of both VPAC1 and VPAC2 receptors in
RT mammals."
RL Endocrinology 0:0-0(1999).
DR EMBL: AF100644; AAD03602.1; -.
DR INTERPRO: IPR000832; -.
DR INTERPRO: IPR001571; -.
DR INTERPRO: IPR001771; -.
DR INTERPRO: IPR001879; -.
DR INTERPRO: IPR002285; -.
DR PFAM: PF00002; 7tm_2; 1.
DR PRINTS: PR00249; GPCRSECRETIN.
DR PRINTS: PR00491; VASOACTIVEIPR.
DR PRINTS: PR01154; VIP1RECEPTOR.
DR PRINTS: PR01156; PACAPRECEPTR.
DR PROSITE: PS00649; G_PROTEIN_REC_F2_1; 1.
DR PROSITE: PS00650; G_PROTEIN_REC_F2_2; 1.
DR PROSITE: PS00650; G_PROTEIN_REC_F2_2; 1.
KW Receptor.
SQ SEQUENCE 444 AA; 50955 MW; 883B25B729314C4C CRC64;

```

```

Query Match 28.0%; Score 814; DB 13; Length 444;
Best Local Similarity 36.8%; Pred. No. 1.3e-62;
Matches 168; Conservative 73; Mismatches 133; Indels 82; Gaps 12;

QY 41 VLKAKVQCELNITAQLOEGE-----GNCFPEWGLICPRGTGKISAVPCPPY 89
Db 17 ILCPVECSINYOELKHEECVNHEDYFNTAVCKRTWDNITCHPSASIGEVVVLQCPGY 76
QY 90 IYDPNH---KGAVFRHNCNPNGTWDFHSLNKTWANYSDCLRFQPDISGKQE-FCERLY 145
Db 77 FSMFTTGTGVNGVSKNCTSEG-WSEMYP-----ATYAAACGFSTNDPTTEQTVFFGAIK 130
QY 146 VMYTVGYSISGSLAVAILIIGYFRRLLHCTRNTHMLHVFVSMFLRAYTSIFVKDRVV---- 201
Db 131 TGYTIGHSLSLSISUTAAIILCIPRKHLCTRNTHMLHVFVSMFLRAYTSIFVKDRVV---- 190
QY 202 ---HAHIGVRELESIMQDDPONSIEATSDVKSOYIGCKIAVVMFIYFLATNYWILVEG 258
Db 191 ESDCHVG-----SVCKAAWVFQYCIANFWILVEG 224
QY 259 LYLNHLIFVAFESDFTKYUWGFILIGWGPAAFAVAAMAVARATLADARCWELSGDIKIY 318
Db 225 LYLNHLIVISFESEKKYFWYILIGWGPSVFITWASLARVYFEDTCWDIESHLWII 284

```

```

QY 319 QAPILAAIGLNFILFNTVRVLTAKIWTETNAVGHDTKQYRKLAQSTLVLYVFGVHYIV 378
Db 285 KTPILVSLVNFILFICIRILVQKLHSPD-VGRNENSOYTRLAKSTLLIPLFGVHYIM 343
QY 379 FVCLPHSFTGLGWEIRHCELFNFSGFFVSIYCYCNGEVAEVKMKMSRWNL-----S 434
Db 344 FAFPPDNFK---VEVKLVFELILGSGFQGVVAVLYCFLNGEVAELKRWRRNLERFMG 400
QY 435 VDMK-RTPPCGS-----RRCGS 450
Db 401 KDMKYHHPSLGSGNGTNEFSTQISMLTKCSPKTRCCS 436

RESULT 7
Q9RIT8
ID Q9RIT8 PRELIMINARY; PRT; 459 AA.
AC Q9RIT8;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE VASOACTIVE INTESTINAL POLYPEPTIDE 1 (VPAC1) RECEPTOR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129SVJ;
RA Hashimoto H., Nishino A., Shintani N., Hagiwara N., Copeland N.G.,
RA Jenkins N.A., Yamamoto K., Matsuda T., Ishihara T., Nagata S.,
RA Baba A.;
RT "Genomic organization and chromosomal location of the mouse vasoactive
RT intestinal polypeptide 1 (VPAC1) receptor."
RL Genomics 58:90-93(1999).
DR EMBL: AB022860; BAA81896.1; -.
DR EMBL: AB022848; BAA81896.1; JOINED.
DR EMBL: AB022849; BAA81896.1; JOINED.
DR EMBL: AB022850; BAA81896.1; JOINED.
DR EMBL: AB022851; BAA81896.1; JOINED.
DR EMBL: AB022852; BAA81896.1; JOINED.
DR EMBL: AB022853; BAA81896.1; JOINED.
DR EMBL: AB022854; BAA81896.1; JOINED.
DR EMBL: AB022855; BAA81896.1; JOINED.
DR EMBL: AB022856; BAA81896.1; JOINED.
DR EMBL: AB022857; BAA81896.1; JOINED.
DR EMBL: AB022858; BAA81896.1; JOINED.
DR EMBL: AB022859; BAA81896.1; JOINED.
DR INTERPRO: IPR000832; -.
DR INTERPRO: IPR001571; -.
DR INTERPRO: IPR001771; -.
DR INTERPRO: IPR001879; -.
DR INTERPRO: IPR002285; -.
DR PFAM: PF00002; 7tm_2; 1.
DR PRINTS: PR00249; GPCRSECRETIN.
DR PRINTS: PR00491; VASOACTIVEIPR.
DR PRINTS: PR01154; VIP1RECEPTOR.
DR PRINTS: PR01156; PACAPRECEPTR.
DR PROSITE: PS00649; G_PROTEIN_REC_F2_1; 1.
DR PROSITE: PS00650; G_PROTEIN_REC_F2_2; 1.
DR PROSITE: PS00650; G_PROTEIN_REC_F2_2; 1.
KW Receptor.
SQ SEQUENCE 459 AA; 52094 MW; C0C3A9AE1ADF611D CRC64;

```

```

Query Match 27.2%; Score 791.5; DB 11; Length 459;
Best Local Similarity 36.6%; Pred. No. 1.2e-60;
Matches 181; Conservative 81; Mismatches 145; Indels 87; Gaps 19;

QY 13 WLMGLSGCLLARA---QLDSGDTIE-----EQIVLVKAKVQCELNITAQLOEGECNCF 64
Db 11 WL-----CVLAGALACALGPAGSRAASPHQCEYLOMIEKORQOCLEE--AQLENKTTGCS 64
QY 65 PEMDGLTCWRPGTGVGKISAVPCPPYIDFN--HKGAVFRHNCNPNGTWDFHSLNKTWANY 122

```

```
QY 76 GTVGKISAVPCPPYIYDFNHKGVAFRHCNPNGTWFDMHSLNKTWANYSDCLRFLOPDISI 135
D 121 GAPGEVAVPCPDYIYDFNHKGVAFRHCNPNGTWFDMHSLNKTWANYSDCLRFLOPDISI 179
QY 136 GKQECERLYVMYTVGYSISFGSLAVAILIIGYFRLHCTRNYIHMLFVFSFMLRATSI 195
D 180 -EREVDFRLGMIYTVGYSVSLASLTAVAILIAYFRLHCTRNYIHMLFVFSFMLRAVSI 238
QY 196 VKDRVVHAHIGVKELES-----IMDDPQNSIEATSDVKSOYIGCKIAVVMFIYELA 248
D 239 VKDAVLYSGATLDEAERLTERELRAIAQAPPPATAA-----GVAGCRVAVTFELYELA 292
QY 249 TNYWILVEGLYHNLNIFVAFSDTKYLMGFILGWGFPAAVFAAWAVARATLADARCW 308
D 293 TNYWILVEGLYHNLNIFVAFSEKYLWGTVEFGWLPVAFVAVWVSVRATLANTGCWD 352
QY 309 LSAGDIKWIYQAPILAAATGLNFIPLNTVRVLATKIWETNAVGHDTKQYRKLAKSTLVL 368
D 353 LSSGNKKWIIQVPIIASIVLNFILFINVRVLATKLRETNAGRCDTROQYRKLAKSTLVL 412
QY 369 VLVFGVHYIVFVCLPHS-FTGLGWEIFRMHCELFNSFGFFVSIYCYCNGEVOAEVKKM 427
D 413 MPLEGVHYIVEMATPYETVSTGLQVQMHYEMLFNSFGFFVSIYCYCNGEVOAEIKKS 472
QY 428 WSRNMLSDVMKRTPPCGRRRCGSLVTVTHSTSSOSQVAAAHAWCLSLAKLPRSPADSLT 487
D 473 WSRMTLALDFRKARSGSSY-SYGPVMSHTSVTVNGVRAGLGLPLSPRLPAAAAATTA 531
QY 488 ATSLYLANGSVTQSTASHTL-----STRNKED-----SGQRDDILMEKPSRPMS 535
D 532 TTNGHPPIPGHT--KPGAPTLPATPAAPKDDGFLNGSCGLDEASAPERPPALLQE 589
QY 536 NPDT 539
D 590 EWET 593

RESULT 4
Q9PVD2 ID Q9PVD2 PRELIMINARY; PRT; 542 AA.
AC Q9PVD2;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE PARATHYROID HORMONE-RELATED PROTEIN RECEPTOR PTH3R.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprinidae; Rasbora; Danio.
RN [1]
RP SEQUENCE FROM N.A.
RA Rubin D.A., Jueppner H.;
RT "Zebrafish Express the Common Parathyroid Hormone/Parathyroid Hormone-
RT related Peptide Receptor (PTHr) and a Novel Receptor (PTH3R) That Is
RT Preferentially Activated by Mammalian and Fuguish Parathyroid
RT Hormone-Related Peptide."
RL J. Biol. Chem. 274:28185-28190(1999).
DR EMBL: AF132085; AAF01266.2; -
DR INTERPRO: IPR000832; -
DR INTERPRO: IPR001879; -
DR INTERPRO: IPR002170; -
DR PFAM: PF00002; 7tm_2; 1.
DR PRINTS: PR00249; GPCRSECRETIN.
DR PRINTS: PR00393; PTHHORMONER.
DR PROSITE: PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE: PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 542 AA; 61438 MW; 08688658E2727303 CRC64;
```

Query Match 45.0%; Score 1307.5; DB 13; Length 542;
Best Local Similarity 48.9%; Pred. No. 2.5e-105;
Matches 260; Conservative 84; Mismatches 123; Indels 65; Gaps 10;

```
QY 19 CLL--ARAQLSDGTITIEQIVLVKAKVOCELNITAOQL-EGEGNCFPBWDGLICWPR 75
D 14 CVLMGARALIDSDVITRDEQIFLLIGARSCEERTIRAOQSDVVRNENCAPEWDGIIICPT 73
QY 76 GTVGKISAVPCPPYIYDFNHKGVAFRHCNPNGTWFDMHSLNKTWANYSDCLRFLOPDISI 135
D 74 GKPQMVAVLCPYIYDFNHGAYRHCDAAGNQEQVSIINRTWANYTECTTYLHTNHS- 132
QY 136 GKQECERLYVMYTVGYSISFGSLAVAILIIGYFRLHCTRNYIHMLFVFSFMLRATSI 195
D 133 DQEEVEERLYLMYTYGYSISLAALLVAVSILCYFKRLHCTRNYIHILFTSFICRAISIF 192
QY 196 VKDRVVHAHIGVKELESIMDDPQNSIEATSDVKSOYIGCKIAVVMFIYFLATNYWIL 255
D 193 VKDAVLYAVTNDGELED-----GAVQRPWVGCKAAVTFELYLATNHYWIL 239
QY 256 VEGLYLHNLIFVAFSDTKYLMGFILGWGFPAAVFAAWAVARATLADARCWELSGADIK 315
D 240 VEGLYLHSLIFMAFLSDKNCLWALTIIIGWIPAVEFSIWSVARSYSLADTQCDWISAGNLK 299
QY 316 WIYQAPILAAATGLNFIPLNTVRVLATKIWETNAVGHDTKQYRKLAKSTLVLVFGVH 375
D 300 WIYQVPIALAAIVVAFELFINIRVLASKLWETNTCKLDPQOYRKLKSTVLMPLFGVH 359
QY 376 YIVFVCLPHS-FTGLGWEIFRMHCELFNSFGFFVSIYCYCNGEVOAEVKKMWRNLS 434
D 360 YMLFMALPYDTVDTGLLQIQMHYEMLFNSGQFFVAFIYCFNGEVOAEVKKWLRLSLA 419
QY 435 VDMKR-----TPPCGRRRCGSLVTVTHSTSSOSQVAAAHAWCLSLAKLPRSPADSLTA 488
D 420 LDLKQKARVHSSACGSGYGYGM----SHITTTQ----- 449
QY 489 TSLYLANGSVTQSTASHTLSTRNKEDSGRQRDDILMEKPSRPMSNPDT 540
D 450 -SVCLSVSGA---KGGHSLHTIGAKGOSHLOHSGNL-----PGYAPQDTE 490
```

```
RESULT 5
Q9R1D4 ID Q9R1D4 PRELIMINARY; PRT; 169 AA.
AC Q9R1D4;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE PARATHYROID HORMONE TYPE-2 RECEPTOR PRECURSOR (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE; 99367425.
RA Rubin D.A., Hellman P., Zon L.I., Lobb C.J., Bergwitz C., Juppner H.;
RT "A G protein-coupled receptor from zebrafish is activated by human
RT parathyroid hormone and not by human or teleost parathyroid hormone-
RT related peptide. Implications for the evolutionary conservation of
RT calcium-regulating peptide hormones."
RL J. Biol. Chem. 274:23035-23042(1999).
DR EMBL: AF132083; AAD51909.1; -
DR INTERPRO: IPR000832; -
DR PFAM: PF00002; 7tm_2; 1.
DR PRINTS: PR00249; GPCRSECRETIN.
KW Receptor.
FT NON_TER 1
FT NON_TER 169
SQ SEQUENCE 169 AA; 19674 MW; 748CC8231F1C69EA CRC64;
```

Query Match 29.5%; Score 858.5; DB 11; Length 169;
Best Local Similarity 92.9%; Pred. No. 5.6e-67;
Matches 158; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

```
Db 115 PCQYIYDFNHAAYRRCDSNGSVLAESSNKTWVNYTECIK--SPEPNKKQVFFERL 172
QY 145 YVMTYVGSISFGSLAVAILIIGYFRRLHCTRNTHMHFLVSPMLRATISFVKDRVVHAH 204
Db 173 HIMYTVGAVSFSSLLVAIFIIGYFRRLHCTRNTHMHFLVSPMLRAASIFVKDHVHTS 232
QY 205 IGKLESLIMODDQNSIATSVDKSOYIGCKTAVVWFYFLATNYWILVEGLYLNH 264
Db 233 AGLESDAVLM--NNFTAVADVADVTQYMGCKVTVLFIYFLATNYWILVEGLYLNH 291
QY 265 IFVAFESDTKYLWGFILIGNGFPAFAVAANAVARATLADARCWELSGADLKIWIQAPL 324
Db 292 IFMAFLSDSKYLWGFILIGNGFPAFAVAANAVARATLADARCWELSGADLKIWIQAPL 351
QY 325 AIGLNFILFNLTVRLVATKIWETNAVGHDRKQYRKLAQSLVVLVFGVHYIVFVCLPH 384
Db 352 AIGLNFILFNLTVRLVATKIWETNAVGHDRKQYRKLAQSLVVLVFGVHYIVFVCLPH 411
QY 385 SFTGLGWEIRHMCBELFNSFGQFVSIYCYNGEVAQAEVKKMWSRNLSVDWKRTPPCG 444
Db 412 TFEGLGWEIRHMCBELFNSFGQFVSIYCYNGEVAQAEVKKMWSRNLSVDWKRTPPCG 471
QY 445 SRRGSLVLTTHSTSSOSQVAAHANCLSLAKLPRSPADSLTATSLYAMSGVTSQRTA 504
Db 472 SNRYGSLVLTGNNSTSSOSQVAAHANCLSLAKLPRSPADSLTATSLYAMSGVTSQRTA 522
QY 505 S--HTLSRNSK-----EDSGRQDDILMEK--PSRPMESNPDE 540
Db 523 TLPQVYVNSDADSLPSPPEPEPSAKQVDDILLKESLTPRSPSGLEDDE 572
```

RESULT 2

```
Q9PVD3 ID Q9PVD3 PRELIMINARY; PRT; 536 AA.
AC Q9PVD3;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE PARATHYROID HORMONE RECEPTOR PTH1R.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprinidae; Rasbora; Danio.
RN [1]
RP SEQUENCE FROM N.A.
RA Rubin D.A., Jueppner H.;
RT "Zebrafish Express the Common Parathyroid Hormone/Parathyroid Hormone-
RT related Peptide Receptor (PTH1R) and a Novel Receptor (PTH3R) That is
RT Preferentially Activated by Mammalian and Fuguish Parathyroid
RT Hormone-related Peptide."
RL J. Biol. Chem. 274:28185-28190 (1999).
DR EMBL: AF132084; AAF01265.1; -.
DR INTERPRO: IPR000832; -.
DR INTERPRO: IPR001879; -.
DR INTERPRO: IPR002170; -.
DR PFAM: PF00002; 7tm_2; 1.
DR PRINTS: PR00249; GPCRSECRETIN.
DR PRINTS: PR00393; PTHORMONER.
DR PROSITE: PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE: PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 536 AA; 61454 MW; 65380A56DB2C8FA9 CRC64;
```

Query Match 48.88; Score 1418.5; DB 13; Length 536;
Best Local Similarity 57.88; Pred. No. 5.8e-115;
Matches 269; Conservative 80; Mismatches 107; Indels 9; Gaps 6;

```
QY 4 LGASLHV--WGLWMLGSL--LARAQLSDGTITIEEQIVLVKAKVOCELNITAQLOGE 60
Db 1 MGATLIVRTGLFCGTLISVYGLVDADDVLTKEEQIYLLFNARKKCRRAIKSKHTSE 60
```

```
QY 61 GNCPEPNDGLICWPRGVTKISAVPCPPYIYDFNHKGAVFRHCNPNGTWDFMHLNKTWA 120
Db 61 GSCLPNDGILCWPEGVPGKMWSTSCPEYIYDFNHKGAYRRCDLNGTWELASHNNKTWA 120
QY 121 NYSCLAFLOPDISIGKQECERLYVMYTCYSISFSGSLAVAILIIGYFRRLHCTRNVIH 180
Db 121 NYSCLAFLOPDISIGKQECERLYVMYTCYSISFSGSLAVAILIIGYFRRLHCTRNVIH 180
QY 181 MHLFVSFMLRATISFVKDRVVHAHIGVKELESLIMODDQNSIATSVDKSOYIGCKTAV 240
Db 181 MHLFVSFMLRATISFVKDRVVHAHIGVKELESLIMODDQNSIATSVDKSOYIGCKTAV 240
QY 241 VMFYFLATNYWILVEGLYLNHILFVAFSDTKYLWGFILIGNGFPAFAVAANAVARAT 300
Db 241 VMFYFLATNYWILVEGLYLNHILFVAFSDTKYLWGFILIGNGFPAFAVAANAVARAT 300
QY 301 LADARCWELSGADLKIWIQAPLAAIGLNFILFNLTVRLVATKIWETNAVGHDRKQYRK 360
Db 301 LADARCWELSGADLKIWIQAPLAAIGLNFILFNLTVRLVATKIWETNAVGHDRKQYRK 360
QY 361 LAKSTLVLVVFGVHYIVFVCLPHS--FTGLGWEIRHMCBELFNSFGQFVSIYCYNGE 419
Db 361 LAKSTLVLVVFGVHYIVFVCLPHS--FTGLGWEIRHMCBELFNSFGQFVSIYCYNGE 419
QY 420 VQAEVKKMWSRNLSVDWKRTPPCGSR--RCGSVL--TTVTHSTS 460
Db 420 VQAEVKKMWSRNLSVDWKRTPPCGSR--RCGSVL--TTVTHSTS 460
```

RESULT 3

```
Q9TU31 ID Q9TU31 PRELIMINARY; PRT; 595 AA.
AC Q9TU31;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE PARATHYROID HORMONE RECEPTOR-1.
GN PTH1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN [1]
RP SEQUENCE FROM N.A.
RA Smock S.L., Vogt G.A., Castleberry T.A., Lu B., Owen T.A.;
RT "Molecular cloning and functional characterization of the canine
RT parathyroid hormone receptor-1 (PTH1)."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF167095; AAD55938.1; -.
DR INTERPRO: IPR000832; -.
DR INTERPRO: IPR001879; -.
DR INTERPRO: IPR002170; -.
DR PFAM: PF00002; 7tm_2; 1.
DR PRINTS: PR00249; GPCRSECRETIN.
DR PRINTS: PR00393; PTHORMONER.
DR PROSITE: PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE: PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW Receptor.
SQ SEQUENCE 595 AA; 66308 MW; 09568ECF38D4D258 CRC64;
```

Query Match 46.3%; Score 1345.5; DB 6; Length 595;
Best Local Similarity 46.5%; Pred. No. 1.4e-108;
Matches 281; Conservative 90; Mismatches 154; Indels 79; Gaps 11;

```
QY 4 LGASLHVWGLWMLGSL---ARAQLSDGTITIEEQIVLVKAKVOCELNITAQLO--- 57
Db 1 MGAVRIAPGLALLCCPVLSSAYALVDADVMTKEEQIFLLHRAQAQCKRLKEVLQSPA 60
QY 58 -----ECEGN-----CFPNDGLICWPR 75
Db 61 DIMESDKWASASTSGPKKKEKASKGLYPESEDEKQVPTGSRHGRPCLPEDWHLICWPL 120
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:47:29 ; Search time 152.43 Seconds
(without alignments)
331.401 Million cell updates/sec

Title: US-09-236-468a-2

Perfect score: 2907

Sequence: 1 MAWLGLSLHVMGLMLGSL.....DDLMEKPSRPMESNPDTEG 541

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 297973

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTREMBL_14.*

2: sp_archaea.*

3: sp_bacteria.*

4: sp_fungi.*

5: sp_human.*

6: sp_invertebrate.*

7: sp_mammal.*

8: sp_mhc.*

9: sp_organelle.*

10: sp_phase.*

11: sp_plant.*

12: sp_rodent.*

13: sp_virus.*

14: sp_vertebrate.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1677.5	57.7	575	13 Q9PWB7	Q9pwb7 brachydanio
2	1418.5	48.8	536	13 Q9PVD3	Q9pvd3 brachydanio
3	1345.5	46.3	595	6 Q9TU31	Q9tu31 canis famil
4	1307.5	45.0	542	13 Q9PVD2	Q9pvd2 brachydanio
5	858.5	29.5	165	11 Q9RLD4	Q9rlid4 mus musculus
6	814	28.0	444	13 Q9YHC6	Q9yhc6 rana ridibu
7	791.5	27.2	459	11 Q9RIT8	Q9rit8 mus musculus
8	713	24.5	465	13 Q93769	Q93769 carassius a
9	702.5	24.2	438	13 Q93768	Q93768 carassius a
10	685.5	23.6	465	13 Q9PDK1	Q9ptk1 xenopus lae
11	665	22.9	550	11 Q9Z0W0	Q9z0w0 rattus norv
12	635.5	21.9	553	4 Q95838	Q95838 homo sapien
13	633	21.8	441	6 Q9TUJ0	Q9tujo bos taurus
14	610	21.0	404	6 Q9TUJ1	Q9tunj1 bos taurus
15	604	20.8	439	11 Q9WU99	Q9wu99 rattus norv
16	592	20.4	492	13 Q9PDK1	Q9ptk1 hoptobarac
17	534.5	18.4	463	11 Q9WU92	Q9wup2 mus musculus
18	534.5	18.4	463	11 Q9QXH8	Q9qxh8 mus musculus
19	525.5	18.1	463	11 Q9RIW5	Q9riw5 mus musculus

20	519.5	17.9	504	5 Q9V716	Q9v716 drosophila
21	505.5	17.4	388	5 Q9V6C7	Q9v6c7 drosophila
22	487	16.8	126	13 Q57671	Q57671 meleagris g
23	462	15.9	401	4 Q9UK64	Q9uk64 homo sapien
24	439	15.1	94	13 Q9P8G1	Q9prg1 lctalurus p
25	425.5	14.6	633	5 Q9W4Y2	Q9w4y2 drosophila
26	418	14.4	168	13 Q9YHC7	Q9yhc7 rana ridibu
27	387	13.3	561	5 Q9VYH9	Q9vyh9 drosophila
28	376	12.9	444	5 Q9V6N5	Q9v6n5 drosophila
29	347.5	12.0	202	13 Q98955	Q98955 meleagris g
30	328.5	11.3	167	13 Q9YHC8	Q9yhc8 rana ridibu
31	322.5	11.1	454	5 Q9U3P6	Q9u3p6 caenorhabdi
32	249.5	8.6	1449	5 Q10922	Q10922 caenorhabdi
33	235	8.1	130	11 Q70440	Q70440 mus musculu
34	233	8.0	1487	11 Q92174	Q92174 rattus norv
35	227	7.8	1515	11 Q88917	Q88917 rattus norv
36	226.5	7.8	1478	11 Q88923	Q88923 rattus norv
37	225	7.7	1467	6 Q97830	Q97830 bos taurus
38	225	7.7	1472	6 Q97831	Q97831 bos taurus
39	224	7.7	1474	4 Q94910	Q94910 homo sapien
40	223	7.7	1283	6 Q97823	Q97823 bos taurus
41	223	7.7	1351	6 Q97829	Q97829 bos taurus
42	223	7.7	1512	6 Q97821	Q97821 bos taurus
43	223	7.7	1580	6 Q97827	Q97827 bos taurus
44	220	7.6	1240	6 Q97822	Q97822 bos taurus
45	220	7.6	1308	6 Q97828	Q97828 bos taurus

ALIGNMENTS

RESULT 1

Q9PWB7

ID Q9PWB7 PRELIMINARY; PRT: 575 AA.

AC Q9PWB7;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

DE PARATHYROID HORMONE TYPE-2 RECEPTOR PRECURSOR.

OS Brachydanio rerio (zebrafish) (zebra danielo).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;

OC Cypriniformes; Cyprinidae; Cyprinidae; Cyprinidae; Danio.

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=KIDNEY;

RX MEDLINE; 99367425.

RA Rubin D.A., Hellman P., Zon L.I., Lobb C.J., Bergwitz C., Juppner H.;

RT "A G protein-coupled receptor from zebrafish is activated by human

parathyroid hormone and not by human or teleost parathyroid hormone-

related peptide. Implications for the evolutionary conservation of

calcium-regulating peptide hormones.";

RL J. Biol. Chem. 274:23035-23042(1999).

DR EMBL; AF132082; AAB51908.1; -.

DR INTERPRO; IPR001879; -.

DR PFAM; PF00002; 7tm.2; 1.

DR PRINTS; PR00249; GPCRSECRETIN.

DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.

DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.

SQ SEQUENCE 575 AA; 64244 MW; 888F1C4DDB3A14DC CRC64;

Query Match

Best Local Similarity 57.7%; Score 1677.5; DB 13; Length 575;

Matches 322; Conservative 72; Mismatches 109; Indels 27; Gaps 7;

Qy 25 QLDSDGTTIEEQIVLVKAKVQCELNTAQLQEGEGNCFEWDGLICWPGTGVKISAV 84

Db 56 QAGEDTEAEQVQMLLDLAKLQCLQKVSSD-DPAVGVCPVEWDGLICWPGTGLTKT 114

Qy 85 PCPPYIVDFNHKGVAFRHCNPNGTWDFMHSLNKTNWYSDCLRFLLQPDISIGKQFCERL 144

01-APR-1993 (Rel. 25, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
VASOACTIVE INTESTINAL POLYPEPTIDE RECEPTOR 1 PRECURSOR (VIP-R-1)
(PITUITARY ADENYLATE CYCLASE ACTIVATING POLYPEPTIDE TYPE II RECEPTOR)
(PACAP TYPE II RECEPTOR) (PACAP-R-2).
VIPRI.
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ON Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1]
RN TISSUE=LUNG.
RC SEQUENCE FROM N.A.
RX MEDLINE; 92232309.
RA Ishihara T., Shigemoto R., Mori K., Takahashi K., Nagata S.;
RT "Functional expression and tissue distribution of a novel receptor
for vasoactive intestinal polypeptide.";
RL Neuron 8:811-819(1992).
CC -!- FUNCTION: THIS IS A RECEPTOR FOR VIP. THE ACTIVITY OF THIS
RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL
CYCLASE.
CC
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC
CC -!- TISSUE SPECIFICITY: IN LIVER, LUNG, INTESTINES, THYMUS AND BRAIN
(MOSTLY IN THE CEREBRAL CORTEX AND HIPPOCAMPUS).
CC
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

CC EMBL; M86835; AAA42331.1; .
CC PIR; JH0594; JH0594.
CC GCRDB; GCR_0369; .
DR DR INTERPRO; IPR000832; .
DR DR INTERPRO; IPR001571; .
DR DR INTERPRO; IPR001771; .
DR PFAM; PF00002; 7tm_2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PRINTS; PR00491; VASOACTVEIPR.
DR PRINTS; PR01154; VIPIRECEPTOR.
DR DR PROSITE; PS00649; G_PROTEIN_RECP_F2_1; 1.
DR DR PROSITE; PS00650; G_PROTEIN_RECP_F2_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 30 POTENTIAL.
FT FT CHAIN 31 459 VASOACTIVE INTESTINAL POLYPEPTIDE
RECEPTOR 1.
FT FT DOMAIN 31 143 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 144 168 1 (POTENTIAL).
FT FT DOMAIN 169 175 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 176 195 2 (POTENTIAL).
FT FT DOMAIN 196 217 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 218 241 3 (POTENTIAL).
FT FT DOMAIN 242 255 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 256 277 4 (POTENTIAL).
FT FT DOMAIN 278 294 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 295 318 5 (POTENTIAL).
FT FT DOMAIN 319 343 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 344 363 6 (POTENTIAL).
FT FT DOMAIN 364 375 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 376 395 7 (POTENTIAL).
FT FT DOMAIN 396 459 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 58 58 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 69 69 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 100 100 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 459 AA; 52057 MW; 99E9B57DA86698D2 CRC64;

Query Match 26.6%; Score 772.5; DB 1; Length 459;
Best Local Similarity 36.8%; Pred. No. 4.5e-52;

Matches 172; Conservative 78; Mismatches 134; Indels 83; Gaps 36

QY 36 EQIVLVILKARVQCLNITAQIQEGNCGNCFPEWDLGICWPRGTGKISAVPCP-----PY 89

Db 38 EYLQIEIQOQCLEE--AQLENETTGCCKMWDNLTCWPTTPRQAVVLDCLPIFOLFAP 95

QY 90 IYDNIHKGVAFRCPNPNGTWDFWHLSLNTKWTANYSDCLRFLQP---DISIG-----136

Db 96 IHGYN-----ISRSTCEG-----WSQ-----LEPGPYHTACGLNDRASSLDE 133

QY 137 --KQECERLYVMVTCYSISFGSLAVAILIGYFRRLHCTRNYIHMHLPVSEMLRATSI 194

Db 134 QQQTKFYNTVKTGTGIGYSUSLASLLVAMAILSEFRKLHCTRNYIHMHLPMSFILRATAV 193

QY 195 FVKDRVVHAHTGVKELESIMQDDPQNSIEATSVDRKSYIGICKIAVVMFYFIATVNYWI 254

Db 194 FIKDMALF-----NSGEIDHCSEAS-VGCKAAVVVFQYCVMANFFWL 234

QY 255 LVEGLYLHNLIFVAFPSDTKYLLNGFPLIGWGFFPAAVVAWAVARATLADARCWE-LSAGD 313

Db 235 LVEGLYLTYLLAVSFFSERKYFGWYILIGWVPSVFETTIVTVRVIRYFDFGCDWTIINS 294

QY 314 IKWIIQAPILAAICLNFILFELNVTVRVLATKIWETNAVGHDRHQYRKLAKSLTLVLVLPFG 373

Db 295 LMWIIKAPILSLILNVFLVFICTIRILVQKLRPDP-IGKNDSSPYSRLLAKSTLLIPLIFG 353

QY 374 VHYTVFVCLPSSHFTGLGWEIRMHCELFNFGFQFVFSIIVCYCNGEVOAEVKKMWSRWNL 433

Db 354 IHYVMAFFPDNKA---QVMMVFELVVGSGFQGVVAILLYCFUNGVEQAEILRRKWRWHL 410

QY 434 S-VDM--KRTPPCGSRR---CG---SVLPTVT-----HSTSSSQVA 466

Db 411 QGVILGSSKSOHPGGSGNGATCTQVSMLTRVSPSARRSSSFQAEVS 457

RESULT 14

VIPIR_HUMAN STANDARD; PRT; 457 AA.

ID AC VIPR_HUMAN Q15871;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE VASOACTIVE INTESTINAL POLYPEPTIDE RECEPTOR 1 PRECURSOR (VIP-R-1)

DE (PITUITARY ADENYLATE CYCLASE ACTIVATING POLYPEPTIDE TYPE II-RECEPTOR)

DE (PACAP TYPE II RECEPTOR) (PACAP-R-2).

GN VIPR1.

OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

[1]

RN RP SEQUENCE FROM N.A.

RC TISSUE=INTESTINE;

RC MEDLINE; 93290641.

RA Sreedharan S.P., Patel D.R., Huang J.-X., Goetzel E.J.;

RT "Cloning and functional expression of a human neuroendocrine

RT vasoactive intestinal peptide receptor."

RL Biochem. Biophys. Res. Commun. 193:546-553(1993).

RN [2]

RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

RC TISSUE=INTESTINE;

RC MEDLINE; 94235025.

RA Couvineau A., Rouyer-Pessard C., Darmoul D., Maoret J.J., Carretero I.,

RA Ogier-Denis E., Laburthe M.;

RT "Human intestinal VIP receptor: cloning and functional expression of

RT two cDNA encoding proteins with different N-terminal domains.";

RL Biochem. Biophys. Res. Commun. 200:769-776(1994).

RN [3]

RP SEQUENCE OF 33-457 FROM N.A.

RC TISSUE=LIVER;

RC MEDLINE; 95001220.

RA Gagnon A.W., Aiyar N., Elshourbagy N.A.;

RT "Molecular cloning and functional characterization of a human liver

RT vasoactive intestinal peptide receptor.";


```
Db 82 KASGKFYDPSKENDKDVPTGSRRRRRCPUPEWDNIVCWPPLGAPGEVAVPCPDYIDFNHK 141
Qy 97 GVAFHCPNCTWDFMHSINKTWNYSOCLRFLQPDISIGKOEFCERLYVMVTCYSISF 156
Db 142 GHAYRCDRNGSWEVPGHNRTWANYSECLAFMTNTR--BREVDFRLGMYTTCYSMSL 199
Qy 157 GSLAVAILIYGFRLLHCTRNYIIMHLFVSPMLRATSFIVKDRVVAHIGVKELES--- 213
Db 200 ASLTAVAVILAYFRLHCTRNYIIMHMFSLFMLRAASFVKAVALYSCFTLDEARLTEE 259
Qy 214 ----IMQDDPNSTSEATSDKSOYIGCKIAVVMFYFLATNYYILVEGLYHNLFIWAF 269
Db 260 ELHIIAQVPPPPAAAVG-----YAGCRVAVTFYFLATNYYILVEGLYHLSLIFMAF 314
Qy 270 FSDYKLMGFTLIGFPFAAFAVAWAVARATLADARCELSAGDTKVIQYAPILAAIGLN 329
Db 315 FSEKYLWGFIFGWLPAVFAVAVGVATLANTGCDLSSGHKKHIIQVPIIASVVLN 374
Qy 330 FILELNTVRVLATKIWETNAVGHDTKQYRKLRKSTLVLVFVGHYIVFVCLPHS-FTG 388
Db 375 FILELNIIRVLATKLRETNACRCDTRQYRKLRLSTLVLPVLFVGHYIVFVMAIYTEVSG 434
Qy 389 LGWEIRMCLEFFNSFGQFFVSIYCYNGEVAEQVKMWSRWNLSDVDMKRTPPCGSRRC 448
Db 435 TLWQTMHYELFNSFGQFFVSIYCYNGEVAEQVKMWSRWNLSDVDMKRTPPCGSRRC 494
Qy 449 GSVLTTVTHSTSSOSOVAHAHWCLSLAKLPRSPADSLTATSLYLAWSG-----VTQS 501
Db 495 -SYGPMVSHTSVTVNGPRAG-----LSLPLSPRLPP---ATTNGHSQPLGHPAKGPATET 546
Qy 502 RTASHTLSTRS-----NKESGRQRDDILMEK 528
Db 547 ETLPVTMAVPKDDGFLNGSCSLDEASGSARPPPLQEQ 585

RESULT 7
ID PTRR_MOUSE STANDARD; PRT; 591 AA.
AC P41593; 062119;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR
DE PRECURSOR (PTH/PTHr RECEPTOR).
GN PTHr1 OR PTHr.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-C3H/HEHA;
RA MEDLINE; 95034305.
RA Karperien M., van Dijk T.B., Hoelijmakers T., Cremers F.,
RA Abou-Samra A.-B., Boonstra J., de Laat S.W., Defize L.H.K.;
RT "Expression pattern of parathyroid hormone/parathyroid hormone
RT related peptide receptor mRNA in mouse postimplantation embryos
RT indicates involvement in multiple developmental processes.";
RL Mech. Dev. 47:29-42(1994).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RA MEDLINE; 94255468.
RA McCuaig K.A., Clarke J.C., White J.H.;
RT "Molecular cloning of the gene encoding the mouse parathyroid
RT hormone/parathyroid hormone-related peptide receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:5051-5055(1994).
CC -1- FUNCTION: THIS IS A RECEPTOR FOR PARATHYROID HORMONE AND FOR
CC PARATHYROID HORMONE-RELATED PEPTIDE. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATES ADENYLYL
CC CYCLASE AND ALSO A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER
CC SYSTEM.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
```

Proc. Natl. Acad. Sci. U.S.A. 89:2732-2736(1992).

[2]

SEQUENCE FROM N.A.

RA Pausova Z., Bourdon J., Clayton D., Mattei M.-G., Seidlin M.F., Janicic N., Riviere M., Szpirer J., Levan G., Szpirer C.; "Cloning of a parathyroid hormone/parathyroid hormone-related peptide receptor (PTHr) cDNA from a rat osteosarcoma (UMR 106) cell line: chromosomal assignment of the gene in the human, mouse, and rat genomes.";

RA Genomics 20:20-26(1994).

CC -!- FUNCTION: THIS IS A RECEPTOR FOR PARATHYROID HORMONE AND FOR PARATHYROID HORMONE-RELATED PEPTIDE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATES ADENYL CYCLASE AND ALSO A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.

CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

CC EMBL: M71184; AAA41811.1; -; EMBL: L19475; AAA68098.1; -; CC RDDB: GCR_0206; -; CC RDDB: GCR_0938; -; CC INTERPRO: IPR000832; -; CC INTERPRO: IPR002170; -; CC PFAM: PF00002; 7tm.2; 1; CC PRINTS: PR00249; GPCRSECRETIN. CC PRINTS: PR00393; PTHORMONER CC PROSITE: PS00649; G_PROTEIN_RECEP_F2_1; 1. CC PROSITE: PS00650; G_PROTEIN_RECEP_F2_2; 1. CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal. CC SIGNAL 1 26 CC CHAIN 27 591

FT DOMAIN 27 188 PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR. FT TRANSMEM 189 212 EXTRACELLULAR (POTENTIAL). FT DOMAIN 214 219 CYTOPLASMIC (POTENTIAL). FT TRANSMEM 220 239 2 (POTENTIAL). FT DOMAIN 240 282 EXTRACELLULAR (POTENTIAL). FT TRANSMEM 283 306 3 (POTENTIAL). FT DOMAIN 307 320 CYTOPLASMIC (POTENTIAL). FT TRANSMEM 321 342 4 (POTENTIAL). FT DOMAIN 343 361 EXTRACELLULAR (POTENTIAL). FT TRANSMEM 362 382 5 (POTENTIAL). FT DOMAIN 383 409 CYTOPLASMIC (POTENTIAL). FT TRANSMEM 410 428 6 (POTENTIAL). FT DOMAIN 429 440 EXTRACELLULAR (POTENTIAL). FT TRANSMEM 441 463 7 (POTENTIAL). FT DOMAIN 464 591 CYTOPLASMIC (POTENTIAL). FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL). FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL). FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL). FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL). CC SEQUENCE 591 AA; 66260 MW; 21944F3051B9E9C1 CRC64;

Query Match 45.8%; Score 1331.5; DB 1; Length 591; Best Local Similarity 46.6%; Pred. No. 8.8e-95; Matches 270; Conservative

Qy 22 ARAQLSDGTTTIEQILVLKAKVOCCE-----LNITAQLQEGN----- 62

Db 22 AYALVDADVFTKEEQIFLLHRAQAQCKLLKEVLHFAANTMSKDGWTPASTSGKPKKE 81

Qy 63 -----CFPEWDGLICPRGTGVKISAVPCPPYIDFNHK 96

Query Match 46.0%; Score 1336.5; DB 1; Length 593; Best Local Similarity 47.4%; Pred. No. 3.6e-95; Matches 276; Conservative 81; Mismatches 134; Indels 91; Gaps 11;

Qy 22 ARAQLSDGTTTIEQILVLKAKVOCENIQAQLO----- 57

Db 22 AYALVDADVFTKEEQIFLLHRAQAQCKRLLEVLQPASIMESDKGWTSASTSGKPKRD 81

Qy 58 -----EGEGN-----CFPEWDGLICWPRGTGCKISAVPCPPYIDFNHK 96

Db 82 KASGLYPESEDEKEAPTSYGRCPLEPDWDHILCPGACGEVAVPCPDYIDFNHK 141

Qy 97 GVAFRHCNPGTWDFMHSLNKTWANSYDGLRFLQPDISCKOEFCEBRLYVMVTGVYSISF 156

Db 142 GHAYRCDRNGSWELVPGHNRWTWANSYECVKFLNETR--EREVFDRLGMIVTVGYSVL 199

Qy 157 GSIAVAILLIGYFRRLLHCTRNYYTHMLFVFSFMLRAVTSIFVKORVVAHAGVKELESLSL---- 213

Db 200 ASLTAVAILAYFRRLLHCTRNYYTHMLFLFSEMLRAVTSIFVKOAVLYSGATLDEAERLTEE 259

Qy 214 ----IMODDPQNSIEATSDVSKSOYICKTAVVMFIVFLATNYWIIWVEGLYLNLIIFVAF 269

Db 260 ELRAIAQAPPPATAAG-----YACGRVAVTFFLFLATNYWIIWVEGLYLNLIIFMAF 314

Qy 270 FSDTKYLMGFIILGWGFFAAVAAVAVARATLADARCWELSGADIKKIYOAPILAAIGLN 329

Db 315 FSEKKYLMGTFVFGWGLPAVFAVAVSVRATLANTCWDLSGNNKKWIIQVPLASIVLN 374

Qy 330 FILELNTVRVLATKIWETNAVGHDTQRKYLAKSTLVLVLFVGVHYIVFVCLPHS-FTG 388

Db 375 FILEINIVRLATKLRETNAGRCDTQQYRKLLSKLVLMLPLFGVHYIVFEMATPYTEVSG 434

Qy 389 LGWEIRMHCELFNSFGQFFVSIYCYCNGEVOAEVKWMSRWNLSDVMKRTPPCGSRRC 448

Db 435 TLMQVQHMYEMLNSFGQFFVSIYCYCNGEVOAEIKKSNRWTLALDFKRAKSGSSSY 494

Qy 449 G-SVLTVTHTSSQSAHAHACWLSLAKLRPSADSLTATSL-VLAMSQVGTQST-AS 505

Db 495 SYGPMVSHTSVTVNGPRVGLG-----LPLSPRLPTATNGHPQLPGHAKPCTPAL 545

Qy 506 HTLSTRS-----NKEDSGRQRDILMEK 528

Db 546 ETLETTPPAMAAPKDGDFLNGSCGLDEEASGERPPALLQE 587

RESULT 6

PTHR_RAT ID PTHR_RAT STANDARD; PRT: 591 AA.

AC P25961;

DT 01-MAY-1992 (Rel. 22, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR

DE PRECURSOR (PTH/PTHR RECEPTOR).

GN PTHR1 OR PTHR.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

RP [1]

RN SEQUENCE FROM N.A.

RC TISSUE=BONE;

RX MEDLINE: 92212903.

RA Abou-Samra A.-B., Jueppner H., Force T., Freeman M.W., Kong X.-F., Schipani E., Urena P., Richards J., Bonventre J.V., Potts J.T. Jr., Kronenberg H.M., Segre G.V.

RA "Expression cloning of a common receptor for parathyroid hormone and parathyroid hormone-related peptide from rat osteoblast-like cells: a single receptor stimulates intracellular accumulation of both cAMP and inositol trisphosphates and increases intracellular free calcium.";

RT RT


```
Db 1 MPWLALPYICWLLRLRCLLVAGALDSDGTTTIEQIVLVNKAQKMOCELNITAQFORGE 60
Qy 61 GNCFFPEWGLICWPRGTGKTSVPCPPYIVDFENIUKGVAFRHCNPNGTWDFMHSLNKTWA 120
Db 61 GNCFFPEWGLICWPRGTAGKTSAMPCCPSIVDFENHKGVAFRHCTPNGTWDFIHGSKNTWA 120
Qy 121 NYSCLRFLOPDIISGKQFCERLYVMYTVGYSISFGSLAVAILIIGYFRRLHCTRNIIH 180
Db 121 NYSDC--FLQPDINGKQEEFENLYITVCGYSISFGSLAVAILIIGYFRRLHCTRNIIH 178
Qy 181 MHLFVSFMLERATSFVKDRVVAHIGVKELESIMQDDPQNSIBATSVDKSOYIGCKTIAV 240
Db 179 LHLFVSFMLERAXSFVKDRVAQAHLGVALQSLVMQGLQNFIFGGPSVDKSOYVGCKTIAV 238
Qy 241 VMFIYFLATNYWILVEGLYLHNLIFVAFSFTKYLWGFILLIGWGFPAFVAANAVARAT 300
Db 239 VMFIYFLATNYWILVEGLYLHNLIFVSFFSDTKYLWGFILLIGWGFPAFVAVANAVARAT 298
Qy 301 LADARCWELSGDJKWIYOAPTLAAIGLNFILNTVRLVATKIWETNAVGHDTKQYRK 360
Db 299 LADTCEWELSGD--RWIYXXPILAAIGLNFILNTVRLVATKIWETNAVGHDRKQYRK 357
Qy 361 LAKSTPLVLVFGVHYIVFVCLPHTSFTGLGWELRHCLEFFNSFGFPVSYIYCNGEV 420
Db 358 LAKSTPLVLVFGVHYIVFICPHSFGSLWELRHCLEFFNSFGFPVSYIYCNGEV 417
Qy 421 QAEYKKMSRNLSLDWKRTPPCGRRCGSLVTITVTHSTSSQSVAAAHANCLSLAKLPR 480
Db 418 QAEYKKTTNRNLSLDWKAPKCGHRYGSLVTITVTHSTSSQSGMPSTRVLVSSKPAK 477
Qy 481 SPA-----DSLATSLSYLAWSGVTSQRTASHTLSTRNKDQSRQRDDILMEKPSRPMESNP 537
Db 478 TACROIDSHVTLPGYVWSSSDQCOPOS---TPETKKGHGRQEDSDSPVGESSRPVFTI 534
Qy 538 DTEG 541
Db 535 DTEG 538

RESULT 3
PTRR_DIDMA STANDARD; PRT; 585 AA.
AC P25107;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PARATHYROID HORMONE/PARATHYROID HORMONE-RELATED PEPTIDE RECEPTOR
DE PRECURSOR (PTH/PTHR RECEPTOR).
GN PTHR.
OS Didelphis marsupialis virginiana (North American opossum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.
RN [1]
RP MEDLINE; 92054592.
RA Jueppner H., Abou-Samra A.-B., Freeman M., Kong X.-F., Schipani E.,
RA Richards J., Kolakowski L.F. Jr., Hock J., Potts J.T. Jr.,
RA Kronenberg H.M., Segre G.V.;
RT "A G protein-linked receptor for parathyroid hormone and parathyroid
RT hormone-related peptide."
RL Science 254:1024-1026(1991).
CC -1- FUNCTION: THIS IS A RECEPTOR FOR PARATHYROID HORMONE AND FOR
CC PARATHYROID HORMONE-RELATED PEPTIDE. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL
CC CLASE AND ALSO A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER
CC SYSTEM.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M74445; AAA30979.1; -.
DR PIR; A39286; A39286.
DR GCRDB; GCR_0204; -.
DR INTERPRO; IPR000832; -.
DR INTERPRO; IPR002170; -.
DR PFAM; PF00002; 7cml_2; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PRINTS; PR00393; PTHRHORMONER.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 26
FT CHAIN 27 585
FT DOMAIN 27 185
FT TRANSMEM 186 209
FT DOMAIN 210 216
FT TRANSMEM 217 236
FT DOMAIN 237 276
FT TRANSMEM 277 300
FT DOMAIN 301 314
FT TRANSMEM 315 336
FT DOMAIN 337 355
FT TRANSMEM 356 376
FT DOMAIN 377 403
FT TRANSMEM 404 422
FT DOMAIN 423 434
FT TRANSMEM 435 457
FT DOMAIN 458 585
FT CARBOHYD 148
FT CARBOHYD 158
FT CARBOHYD 163
FT CARBOHYD 173
FT SEQUENCE 585 AA; 65963 MW; 34900384CD6DF477 CRC64;
Qy 24 AQLSDGTTITTEQIVLVKAKVOCELNIT-----
Db 24 ALVDADDVITKKEQIILLRNAQACEQRLKEVLRYPELAESAQKDWMSRSKTKKEKPAEK 83
Qy 54 --AQLQEG-----EGNCFPEWDGLICWPRGTGKTSVPCPPYIVDFNKHGVAFR 101
Db 84 LYSQAESREVSRSRQDQGLFPEWONIVCPAGVPGKVAVPCPDYIYDFNKHGRAIR 143
Qy 102 HCNPNGTWDFMHSLNKTWANYSDCLRFLOPDISICKQFCERLYVMYTVGYSISFGSLAV 161
Db 144 RCDNSGWSLWELVPGNNRTWANYSECVKFLTNETR--EREVDFRLGMITYGVISIGSLTV 201
Qy 162 AILIGFPRRLHCTRNIIHMHILFVSFMLERATSFVKDRVVAHIGVKELESIMQDDPQN 221
Db 202 AVLILGVFRLHCTRNIIHMHILFVSFMLERAVSIFTKDAVLYSGVSTDEIER-ITEELRA 260
Qy 222 SIATSVDKSOYIGCKTIAVMFIYFLATNYWILVEGLYLHNLIFVAFSFTKYLWGFILL 281
Db 261 FTEPPPADKAGFVGCGRVAVTFYFLFTNYWILVEGLYLHSLIFMAFFSEKKYLMWGFTL 320
Qy 282 IGWGFPAFVAANAVARATLADARCWELSGDJKWIYOAPTLAAIGLNFILNTVRLVLA 341
Db 321 FGWGLPAFVAVVTVTRATLANTECWDLSSCNKKWILQVPIAAIVNFIINIRVLA 380
Qy 342 TKIWTENAVGHDTKQYRKRLAKSTPLVLVFGVHYIVFVCLPHTS--FTGLGWELRHCLEF 400
Db 381 TKLRETNAGRCDTQQYRKLLKSTPLVLMPLFGVHYIVFMATPYTEVSGILQVQMHYEML 440

Query Match 48.1%; Score 1397; DB 1; Length 585;
Best Local Similarity 48.0%; Pred. No. 8.3e-100;
Matches 275; Conservative 94; Mismatches 134; Indels 70; Gaps 10;

Qy 24 AQLSDGTTITTEQIVLVKAKVOCELNIT-----
Db 24 ALVDADDVITKKEQIILLRNAQACEQRLKEVLRYPELAESAQKDWMSRSKTKKEKPAEK 83
Qy 54 --AQLQEG-----EGNCFPEWDGLICWPRGTGKTSVPCPPYIVDFNKHGVAFR 101
Db 84 LYSQAESREVSRSRQDQGLFPEWONIVCPAGVPGKVAVPCPDYIYDFNKHGRAIR 143
Qy 102 HCNPNGTWDFMHSLNKTWANYSDCLRFLOPDISICKQFCERLYVMYTVGYSISFGSLAV 161
Db 144 RCDNSGWSLWELVPGNNRTWANYSECVKFLTNETR--EREVDFRLGMITYGVISIGSLTV 201
Qy 162 AILIGFPRRLHCTRNIIHMHILFVSFMLERATSFVKDRVVAHIGVKELESIMQDDPQN 221
Db 202 AVLILGVFRLHCTRNIIHMHILFVSFMLERAVSIFTKDAVLYSGVSTDEIER-ITEELRA 260
Qy 222 SIATSVDKSOYIGCKTIAVMFIYFLATNYWILVEGLYLHNLIFVAFSFTKYLWGFILL 281
Db 261 FTEPPPADKAGFVGCGRVAVTFYFLFTNYWILVEGLYLHSLIFMAFFSEKKYLMWGFTL 320
Qy 282 IGWGFPAFVAANAVARATLADARCWELSGDJKWIYOAPTLAAIGLNFILNTVRLVLA 341
Db 321 FGWGLPAFVAVVTVTRATLANTECWDLSSCNKKWILQVPIAAIVNFIINIRVLA 380
Qy 342 TKIWTENAVGHDTKQYRKRLAKSTPLVLVFGVHYIVFVCLPHTS--FTGLGWELRHCLEF 400
Db 381 TKLRETNAGRCDTQQYRKLLKSTPLVLMPLFGVHYIVFMATPYTEVSGILQVQMHYEML 440
```

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 550
FT DOMAIN 27 145
FT TRANSSEM 146 169
FT DOMAIN 170 176
FT TRANSSEM 177 196
FT DOMAIN 197 237
FT TRANSSEM 238 260
FT DOMAIN 261 275
FT TRANSSEM 276 297
FT DOMAIN 298 316
FT TRANSSEM 317 337
FT DOMAIN 338 364
FT TRANSSEM 365 383
FT DOMAIN 384 394
FT TRANSSEM 395 417
FT DOMAIN 418 550
FT CARBOHYD 51 51
FT CARBOHYD 106 106
FT CARBOHYD 116 116
FT CARBOHYD 121 121
SQ SEQUENCE 550 AA; 62235 MW; 2ADD14DBA68A9BF8 CRC64;

Query Match 90.6%; Score 2635; DB 1; Length 550;
Best Local Similarity 91.4%; Pred. No. 1.1e-194;
Matches 500; Conservative 6; Mismatches 29; Indels 12; Gaps 3;

QY 1 MAMLGASHVWGMGLGSCLLARQLSDGTITIEEQIVLVKAKVQCELNITLAQEGE 60
DB 1 MAMLGASHVWGMGLGSCLLARQLSDGTITIEEQIVLVKAKVQCELNITLAQEGE 60
QY 61 GNCPEPDGLICPRGTGKISAVPCPPYIYDFNHKGVAFRHCPNGTWFDMHSLNKTWA 120
DB 61 GNCPEPDGLICPRGTGKISAVPCPPYIYDFNHKGVAFRHCPNGTWFDMHSLNKTWA 120
QY 121 NYSCLRFLOPDISIGKQECERLYVMYTVGYSTFSGLAVAILIIGYFRRLHCTRYIH 180
DB 121 NYSCLRFLOPDISIGKQECERLYVMYTVGYSTFSGLAVAILIIGYFRRLHCTRYIH 180
QY 181 MHLFVSFMLRATSIYKDRVNVHAHIGVKELESIMQDDPQNSIEATSVKDSQYIGCKIAV 240
DB 181 MHLFVSFMLRATSIYKDRVNVHAHIGVKELESIMQDDPQNSIEATSVKDSQYIGCKIAV 240
QY 241 VMFTYFLATNYWLVGLYLHNLIFVAFSTDYKYLIGMGFPAAVFAAWAVARAT 300
DB 241 VMFTYFLATNYWLVGLYLHNLIFVAFSTDYKYLIGMGFPAAVFAAWAVARAT 300
QY 301 LADARCWELSGADIKWYIYQAPILAAIGLNTFLVLTNVTATKIWTETNAVGHDTFRKYRK 360
DB 301 LADARCWELSGADIKWYIYQAPILAAIGLNTFLVLTNVTATKIWTETNAVGHDTFRKYRK 360
QY 361 LAKSTLVLVFVGHYIVFVCLPHSFGLGWEIRHCELFNSFGQFVSIYCYNGEV 420
DB 361 LAKSTLVLVFVGHYIVFVCLPHSFGLGWEIRHCELFNSFGQFVSIYCYNGEV 420
QY 421 QAEVKKMSRWNLSDVKRTPPCGSRRCGSLVLTVTHTSTSSQSOVAAHAWCL---SLAK 477
DB 421 QAEVKKMSRWNLSDVKRTPPCGSRRCGSLVLTVTHTSTSSQSOVAAHAWCL---SLAK 477
QY 478 LPRSPADSLTATSIYLAWSGVQTSRTASHLS---TRSNKEDSGRQRDDILMEKPSRPM 534
DB 478 LPRSPADSLTATSIYLAWSGVQTSRTASHLS---TRSNKEDSGRQRDDILMEKPSRPM 534
QY 481 IASRQPDSD---HITLPGVWVNSBQDCLPHSFHEETKEDSGRQDDILMEKPSRPM 534
DB 481 IASRQPDSD---HITLPGVWVNSBQDCLPHSFHEETKEDSGRQDDILMEKPSRPM 534
QY 535 SNPDTEG 541
DB 535 SNPDTEG 541

RESULT 2

PTH2_RAT

_ID_PTH2_RAT

STANDARD;

PRT;

546 AA.

Query Match 76.1%; Score 2213.5; DB 1; Length 546;

Best Local Similarity 76.5%; Pred. No. 2.1e-162;

Matches 416; Conservative 37; Mismatches 82; Indels 9; Gaps 4;

QY 1 MAMLGASHVWGMGLGSCLLARQLSDGTITIEEQIVLVKAKVQCELNITLAQEGE 60

AC P70555;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE PARATHYROID HORMONE RECEPTOR PRECURSOR (PTH2 RECEPTOR).
GN PTH2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE: 96426194.
RA Usdin T.B., Bonner T.I., Harta G., Mezey E.;
RT "Distribution of parathyroid hormone-2 receptor messenger ribonucleic
RT acid in rat.";
RL Endocrinology 137:4285-4297(1996).
CC -1- FUNCTION: THIS IS A SPECIFIC RECEPTOR FOR PARATHYROID HORMONE. THE
CC ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYL CYCLASE. PTH2 MAY BE RESPONSIBLE FOR PTH EFFECTS IN A
CC NUMBER OF PHYSIOLOGICAL SYSTEMS. IT MAY PLAY A SIGNIFICANT ROLE IN
CC PANCREATIC FUNCTION. PTH2 PRESENCE IN NEURONS INDICATES THAT IT
CC MAY FUNCTION AS A NEUROTRANSMITTER RECEPTOR.
CC -1- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN BRAIN, ARTERIAL AND
CC CARDIAC ENDOTHELIUM. FOUND AS WELL IN SPERM, IN THE HEAD OF THE
CC EPIDIDYMIS. LOWER EXPRESSION IS FOUND IN VASCULAR SMOOTH MUSCLE,
CC EXOCRINE PANCREAS, TESTIS AND PLACENTA.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U55836; AAC52849.1;
DR GCRDB: GCR_1413;
DR INTERPRO: IPR000832;
DR PFAM: PF00002; 7tm.2; 1.
DR PRINTS: PR00249; GPCRSECRETIN
DR PROSITE: PS00649; G-PROTEIN_RECEP_F2_1; 1.
DR PROSITE: PS00650; G-PROTEIN_RECEP_F2_2; FALSE_NEG.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 546
FT DOMAIN 27 143
FT TRANSSEM 144 167
FT DOMAIN 168 174
FT TRANSSEM 175 194
FT DOMAIN 195 235
FT TRANSSEM 236 258
FT DOMAIN 259 273
FT TRANSSEM 274 295
FT DOMAIN 296 313
FT TRANSSEM 314 334
FT DOMAIN 335 361
FT TRANSSEM 362 380
FT DOMAIN 381 391
FT TRANSSEM 392 414
FT DOMAIN 415 546
FT CARBOHYD 51 51
FT CARBOHYD 106 106
FT CARBOHYD 116 116
FT CARBOHYD 121 121
SQ SEQUENCE 546 AA; 61800 MW; 2825AE4040313527 CRC64;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:49:30 ; Search time 58.45 seconds
(without alignments)
295.703 Million cell updates/sec

Title: US-09-236-468A-2

Perfect score: 2907

Sequence: 1 MAWLGLSLHVGWGLMGSL.....DDILMEKPSRPMSNPDTG 541

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 87993

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2635	90.6	550	1	PTR2_HUMAN
2	2213.5	76.1	546	1	PTH2_RAT
3	1397	48.1	585	1	PTRR_DIDMA
4	1345	46.3	585	1	PTRR_PIG
5	1336.5	46.0	593	1	PTRR_HUMAN
6	1331.5	45.8	591	1	PTRR_RAT
7	1324	45.5	591	1	PTRR_MOUSE
8	801	27.6	449	1	SCRC_RAT
9	794	27.3	445	1	SCRC_RABIT
10	785	27.0	458	1	VIPR_PIG
11	784	27.0	440	1	SCRC_HUMAN
12	782	26.9	447	1	VIPR_CARAU
13	772.5	26.6	459	1	VIPR_RAT
14	759.5	26.1	457	1	VIPR_HUMAN
15	713.5	24.5	437	1	VIPS_MOUSE
16	708	24.4	437	1	VIPS_RAT
17	702	24.1	438	1	VIPS_HUMAN
18	685	23.6	468	1	PACR_HUMAN
19	677.5	23.3	496	1	PACR_MOUSE
20	673.5	23.2	513	1	PACR_BOVIN
21	667.5	23.0	523	1	PACR_RAT
22	663	22.8	489	1	GLPR_MOUSE
23	662	22.8	462	1	GLPR_MESAU
24	655	22.5	463	1	GLPR_RAT
25	641.5	22.1	466	1	GLPR_HUMAN
26	630.5	21.7	423	1	GRFR_HUMAN
27	628.5	21.6	463	1	GLPR_HUMAN
28	622.5	21.4	423	1	GRFR_PIG
29	619	21.3	477	1	GLR_HUMAN
30	616	21.2	455	1	GLPR_RAT
31	607	20.9	485	1	GLR_RAT
32	596.5	20.5	423	1	GRFR_MOUSE
33	595	20.5	485	1	GLR_MOUSE

RESULT 1

ID	PTR2_HUMAN	STANDARD;	PRT;	550 AA.
AC	P49190;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	PARATHYROID HORMONE RECEPTOR PRECURSOR (PTH2 RECEPTOR).			
GN	PTH2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.			
RC	TISSUE-BRAIN;			
RX	MEDLINE; 95318121.			
RA	Usdin T.B., Gruber C., Bonner T.I.;			
RT	"Identification and functional expression of a receptor selectively recognizing parathyroid hormone, the PTH2 receptor.";			
RL	J. Biol. Chem. 270:15455-15458(1995).			
RN	[2]			
RP	SEQUENCE OF 26-40 AND 306-550 FROM N.A.			
RX	MEDLINE; 97079671.			
RA	Usdin T.B., Modi W., Bonner T.I.;			
RT	"Assignment of the human PTH2 receptor gene (PTH2) to chromosome 2q33 by fluorescence in situ hybridization.";			
RL	Genomics 37:140-141(1996).			
CC	-I- FUNCTION: THIS IS A SPECIFIC RECEPTOR FOR PARATHYROID HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYL CYCLASE.			
CC	-I- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN BRAIN AND PANCREAS.			
CC	ALSO EXPRESSED IN THE TESTIS.			
CC	-I- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; U25128; AAC50157.1; -			
DR	EMBL; U47124; AAA96796.1; -			
DR	EMBL; U47129; AAC50767.1; -			
DR	EMBL; U47125; AAC50767.1; JOINED.			
DR	EMBL; U47126; AAC50767.1; JOINED.			
DR	EMBL; U47127; AAC50767.1; JOINED.			
DR	EMBL; U47128; AAC50767.1; JOINED.			
DR	GCRDB; GCR_2003; -			
DR	MTM; 601469; -			
DR	INTERPRO; IPR000832; -			
DR	PFAM; PF00002; 7tm_2; 1.			
DR	PRINTS; PR00249; GPCRSECRETIN.			
DR	PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.			
DR	PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.			

34	594.5	20.5	478	1	CALR_CAVPO	00893	cavia porce
35	589.5	20.3	515	1	CALR_MOUSE	Q60755	mus musculus
36	581	20.0	516	1	CALR_RAT	P32214	rattus norv
37	579.5	19.9	490	1	CALR_HUMAN	P30988	homo sapien
38	576	19.8	474	1	CALR_RABIT	P79222	oryctolagus
39	574	19.7	464	1	GRFR_RAT	Q02644	rattus norv
40	554.5	19.1	420	1	CRFR_CHICK	Q90812	gallus gall
41	544.5	18.7	498	1	CALR_PIG	P25117	sus scrofa
42	541	18.6	431	1	CRF2_MOUSE	O60748	mus musculus
43	540.5	18.6	464	1	CGRR_RAT	O63118	rattus norv
44	538.5	18.5	260	1	VIPR_MELGA	Q91085	meleagris g
45	536	18.4	415	1	CRFR_RAT	P35353	rattus norv

ALIGNMENTS

F:206-228/Domain: transmembrane #status predicted <TM3>
F:241-262/Domain: transmembrane #status predicted <TM4>
F:282-304/Domain: transmembrane #status predicted <TM5>
F:329-349/Domain: transmembrane #status predicted <TM6>
F:361-384/Domain: transmembrane #status predicted <TM7>
F:58,88,92/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 24.1%; Score 702; DB 2; Length 438;
Best Local Similarity 36.1%; Pred. No. 2e-51;
Matches 150; Conservative 84; Mismatches 138; Indels 44; Gaps 14;

QY 41 VLKAKVOCENITAQLEGEGNCPEPNDGLICWPRGTGKISAVPCPPYIDFNHK-GVA 99
Db 31 IQEETKCAELLRSSOTKHKR-CSGVNDNITCWPANNGETVTPCPKVSNEFSKAGNI 89

QY 100 FRHCNPNCTDFMHSNLTNTWANTSDCLRFQPDISTIGKQFCELRVYVTVGYISIFGSL 159
Db 90 SKNCTSDG-W-----SETPDFVDACGYSDEPE-DESKITFIYILVKAITYLIGYSVLSMSL 141

QY 160 AVAILIIGYFRRLHCTRYNIHMLHFLVSMRATSIYKDRVVAHIGVKELESIMQDDP 219
Db 142 ATGSIILCLERKLCTRYNIHMLHFLVSMRATSIYKDRVVAHIGVKELESIMQDDP 219

QY 220 QNSTEATSVKSOYQICKIAVVMFYIPLATNYIWLVEGLYHNLIFVAFSDTKYLWGF 279
Db 197 -----SSWVGCKLSLFLQYICIMANFFWLLVEGLYHNLIFVAFSDTKYLWGF 279

QY 280 ILIGWGFPAFAVAAMAVARATLADARCWELSGADIK-WIYQAPITLAAIGNFIPLNTVR 338
Db 245 LLIGWGLPTVCIGAWTAARLYLEDTGCWDTNDHSPVWVIRIPILISIIYVLFISIR 304

QY 339 VLATKIWETNAVGHDTKQYRKLAISTLVLVFVGVHYIVFVCLPHSGFTGLGWEIRMHCE 398
Db 305 ILLQKLTSPDVGND-QSQYKRLAKSTLLIPLFGVHYVFAVFPISISS---KYQILFE 360

QY 399 LFFNSFOGFFVSIYCYCNGEVOAEVKKMSRWNLSDVKRTP-PCGSRR---CGS 450
Db 361 LCLGSGOGLVAVLYCLFNLSEVOCELKRWK-----SRCPTPSASRDYRVCGS 408

RESULT 15
JN0616
Pituitary adenylate cyclase-activating polypeptide receptor 46-5 precursor - rat
N;Alternate names: PACAP receptor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 08-Oct-1999
C;Accession: JN0616; S36768
R;Hosoya, M.; Onda, H.; Ogi, K.; Masuda, Y.; Miyamoto, Y.; Ohtaki, T.; Okazaki, H.; Arima
Biochem. Biophys. Res. Commun. 194, 133-143, 1993
A;Title: Molecular cloning and functional expression of rat cDNAs encoding the receptor
A;Reference number: JN0616; MUID:93326107
A;Accession: JN0616
A;Molecule type: mRNA
A;Residues: 1-467 <HOS>
A;Experimental source: brain
R;Spengler, D.; Waeber, C.; Pantaloni, C.; Holsboer, F.; Bockaert, J.; Seeburg, P.H.; Jock
Nature 365, 170-175, 1993
A;Title: Differential signal transduction by five splice variants of the PACAP receptor.
A;Reference number: S36768; MUID:93382505
A;Accession: S36768
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-467 <SPE>
A;Cross-references: EMBL:Z23279; NID:g404252; PIDN:CAA80817.1; PID:g404253
C;Superfamily: glucagon receptor
C;Keywords: alternative splicing; glycoprotein; receptor
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-467/Product: pituitary adenylate cyclase-activating polypeptide receptor 46-5 #stat
F:47,59,116,299,342,374/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 24.0%; Score 696.5; DB 2; Length 467;

Best Local Similarity 32.4%; Pred. No. 6.1e-51;
Matches 150; Conservative 89; Mismatches 149; Indels 75; Gaps 12;

QY 56 LOEGEGNCPEPNDGLICWPRGTGKISAVPCPPYIDFN----- 94
Db 46 LNESSPCGPGWNDITCWPANNGETVTPCPKVSNEFSKAGNI 89

QY 95 ---HKGVAFRHCNPNCTDFMHSNLTNTWANTSDCLRFQPDISTIGKQFCE-RLYVMTV 150
Db 106 EITDMGVVGRNCTBDG-W-----SEPPHYFDACGDDYEPESGDQDYYVLSVKALYTV 158

QY 151 GYSISFCSLAVAILIIGYFRRLHCTRYNIHMLHFLVSMRATSIYKDRVVAHIGVKELESIMQDDP 219
Db 159 GYSTSLATLTAMVILCFRKLHCTRYNIHMLHFLVSMRATSIYKDRVVAHIGVKELESIMQDDP 219

QY 211 ESLIMQDDPONSIEATSVKSOYQICKIAVVMFYIPLATNYIWLVEGLYHNLIFVAFSDTKYLWGF 279
Db 212 -----EQDSSHCFFYST-----VECKAVMVFHYCVVSNFYFWLFIETGLYFTLLVETFF 259

QY 271 SDTKYLNGFILIGWGFPAFAVAAMAVARATLADARCWELSGADIK-WIYQAPITLAAIGNFIPLNTVR 338
Db 260 PERRYFYWYTIIGWGTPTVCVTWAVLRYFDACGDDYEPESGDQDYYVLSVKALYTV 158

QY 330 FILELNTVRVLATKIWETNAVGHDTKQYRKLAISTLVLVFVGVHYIVFVCLPHSGFTGL 389
Db 320 FVLFIGIILVOKLQSPDMGGNES-SIYLRARSTLLIPLFGIHYTVFAFSPEN---V 375

QY 390 GWEIRMHCELFNSFOGFFVSIYCYCNGEVOAEVKKMSRWNLSDVKRTP-PCGSRR---CGS 450
Db 376 SKRERLVFELGSGFQGVAVLYCLFNLSEVOCELKRWK-----SRCPTPSASRDYRVCGS 408

QY 444 GSRRCGSLTATVTHSTSSQVAAAHAWCLSLAKLPRSPADSL 486
Db 436 SSGVNGCTQLSILSKSSQLRMSSL-----PADNL 465

Search completed: November 8, 2000, 08:53:19
Job time: 350 sec

A: Cross-references: EMBL:Z25895; NID:g414188; PIDN:CAA81104.1; PID:g414189
C: Superfamily: glucagon receptor
C: Keywords: G protein-coupled receptor; intestine; transmembrane protein

Query Match 24.4%; Score 708; DB 2; Length 437;
Best Local Similarity 33.9%; Pred. No. 6.1e-52;
Matches 159; Conservative 89; Mismatches 155; Indels 66; Gaps 15;

QY 41 VLKAKVCELNITAAQLEGEGNCFPEWDGLICPRGTGKISAVPCPPYIYDF-NHKGVA 99
DB 30 IQEETKCAELSSQM-ENHRACSGWDNITCRPADIGETVTVPCPKVFSNYSRPGNI 88
QY 100 FRHCNPNMGTFDWMHSLNKTWANKYSDCLRFQPDISIGQKFCERLYVMYTVGYSISFGSL 159
DB 89 SKNCTSDG-W-----SETPEPDFIDACGYNDPE-DESKITFYILVKAIYTLGYSVSLMSL 140
QY 160 AVALLIIGYFRLHCTRNYTHMILFVSGFMRLRATSFVVKDRVVAHIGVKELESIMODDP 219
DB 141 TTGSIITCLFRKLHCTRNYTHLNLFLSFLRAISVLKDSVLYSSSG-----TLRCHDQP 195
QY 220 QNSIEATSVDKSOYIGCKIAVWFETPLATNYWYLVEGLYLNHLIFVAFSDTKYLLWGF 279
DB 196 -----GSMVGCKLSLVFFQYCI MANFYMLVEGLYLHTLL-VAILPPSRCFLAY 243
QY 280 ILIGWGPFAAFAAWAVARATLADARCWELSSAGDIK-WIYQAPILAAIGLNFILFLNTVR 338
DB 244 LLIGWGPSVCIGAWINTRLSLEDTGCMDTNDHSIPMWIRMPILISIVNFEALFISIVR 303
QY 339 VLATKIWETNAVGHDRKQYRKLAKSTPLVLVFGVHYIVFVCLPHSFTGLGWEIRHCE 398
DB 304 LLQKLTSPDVGND-QSQYKRLAKSTPLLLPLFGVHYMVFAAFP---IGISSTYQILFE 359
QY 399 LFNPSFGFFVSIILCYCNCEVQAEYKMKMSRWNLSDVHWKRTPPCGSRRCGSLVTTVTHS 458
DB 360 LCVGSGFGLVAVLYCFNLSEVQCELRKRRW-----GLCL----- 394
QY 459 TSSQSOVAAAHAACLSLAKLPRPADSLTATSLVYLAMSGVTQSRATSH 507
DB 395 TQGSRDYRLHSMS-----RNGSES-----ALQIHRSRTQSFLOSET 434

submitted to the EMBL Data Library, January 1996
A: Reference number: H01736
A: Accession: G02822
A: Status: preliminary; translated from GB/EMBL/DDBJ
A: Molecule type: mRNA
A: Residues: 1-438 <LUT>
R: Srobooda, M.; Tastenoy, M.; Van Rampelbergh, J.; Goossens, J.F.; De Neef, P.; Waelbr
Biochem. Biophys. Res. Commun. 205, 1617-1624, 1994
A: Title: Molecular cloning and functional characterization of a human VIP receptor fr
A: Reference number: JC2463; MUID: 95110300
A: Accession: JC2463
A: Molecule type: mRNA
A: Residues: 1-38, 'T', '40-411', 'H', '413-423', 'A', '425-438 <SVO>
A: Cross-references: GB:I36566; NID:g550477; PIDN:AAC37569.1; PID:g550478
A: Experimental source: SUP-T1 lymphoblast cell line
C: Genetics:
A: Gene: GDB:VIPR2
A: Cross-references: GDB:335025
A: Map position: 7q36.3-7q36.3
C: Superfamily: glucagon receptor
C: Keywords: glycoprotein; intestine; receptor; transmembrane protein
F: I30-150/Domain: transmembrane #status predicted <TM1>
F: I60-170/Domain: transmembrane #status predicted <TM2>

RESULT 14
G02822
vasoactive intestinal peptide receptor 2 - human
C: Species: Homo sapiens (man)
C: Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 20-Jun-2000
C: Accession: G02822; JC2463
R: Lutz, E.M.
submitted to the EMBL Data Library, January 1996
A: Reference number: H01736
A: Accession: G02822
A: Status: preliminary; translated from GB/EMBL/DDBJ
A: Molecule type: mRNA
A: Residues: 1-438 <LUT>
R: Srobooda, M.; Tastenoy, M.; Van Rampelbergh, J.; Goossens, J.F.; De Neef, P.; Waelbr
Biochem. Biophys. Res. Commun. 205, 1617-1624, 1994
A: Title: Molecular cloning and functional characterization of a human VIP receptor fr
A: Reference number: JC2463; MUID: 95110300
A: Accession: JC2463
A: Molecule type: mRNA
A: Residues: 1-38, 'T', '40-411', 'H', '413-423', 'A', '425-438 <SVO>
A: Cross-references: GB:I36566; NID:g550477; PIDN:AAC37569.1; PID:g550478
A: Experimental source: SUP-T1 lymphoblast cell line
C: Genetics:
A: Gene: GDB:VIPR2
A: Cross-references: GDB:335025
A: Map position: 7q36.3-7q36.3
C: Superfamily: glucagon receptor
C: Keywords: glycoprotein; intestine; receptor; transmembrane protein
F: I30-150/Domain: transmembrane #status predicted <TM1>
F: I60-170/Domain: transmembrane #status predicted <TM2>

Biochem. Biophys. Res. Commun. 206, 246-252, 1995
A:Title: Highly conserved aspartate 68, tryptophane 73 and glycine 109 in the N-terminal
A:Reference number: PC2289; MUID:95118345
A:Accession: PC2289
A:Molecule type: mRNA
A:Residues: 63-129 <CO2>
C:Genetics:

A:Gene: GDB:VIPRI; RCD1; HVR1
A:Cross-references: GDB:128589; OMIM:192321
A:Map position: 3p22-3p22
C:Superfamily: glucagon receptor
C:Keywords: glycoprotein; hormone receptor; intestine; phosphoprotein; transmembrane protein
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-460/Product: vasoactive intestinal peptide receptor #status predicted <MAT>
F:145-168/Domain: transmembrane #status predicted <TM1>
F:176-194/Domain: transmembrane #status predicted <TM2>
F:216-234/Domain: transmembrane #status predicted <TM3>
F:255-277/Domain: transmembrane #status predicted <TM4>
F:299-319/Domain: transmembrane #status predicted <TM5>
F:346-363/Domain: transmembrane #status predicted <TM6>
F:377-396/Domain: transmembrane #status predicted <TM7>
F:58,69,100,293/Binding site: carboxylate (Asn) (covalent) #status predicted
F:76/Binding site: phosphate (Thr) (covalent) (by protein kinase A) #status predicted
F:250/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
F:450/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 26.0%; Score 756; DB 2; Length 460;
Best Local Similarity 36.1%; Pred. No. 5.9e-56;
Matches 180; Conservative 80; Mismatches 144; Indels 94; Gaps 20;

QY 13 WLMGSCLLARAQSDSG-----TITIEQIVLVKAKVQCELNI-TAQLEGEGNCFPE 66
DB 11 WL-----CVLAGALAWALPAGGQAARLQEECDYVQMIQKQLEEAQLENETIGCSKM 66
QY 67 WDGLICHPRTGTVKISAVPCPPYIYDNH---KGVAFPHCNPTWDFMHSLNKTNWYNS 123
DB 67 WDNLTCPATPRGQVVLACPLIKFLFSSTQGRNVS-RSCTDEG-----WTH-- 112
QY 124 DCLRFLOPD-----ISIGKOE--FCERLYVMYTVGYSISFGLSAVALIIGYF 169
DB 113 -----LEPGPYPTACGLDDKAASLDEQOTWFYGSVKTYTIGYGLSLATLLVATILSLF 167
QY 170 RLHCTRNYIHHMLFVSFMLRATSFIVKDRVVAHIGVKELESIMQDDPONSTEATPSVD 229
DB 168 RKLHCTRNYIHHMLFVSFMLRATSFIVKDRVVAHIGVKELESIMQDDPONSTEATPSVD 229
QY 230 KSOYIGCKIAVVMFIYFLATNYWILVEGLYLHNLIFVAFPSDTKYLWGFTILGWGPAA 289
DB 210 EGS-VGCKAAWVFQYCVMANFWLLVEGLYLYTLAVSFPSERKYPFWGYLILGWVPST 268
QY 290 FVAAWAVARATLAD---ARCWELSGADIKTYQAPILAAIGLNFILFNTVRLVATKIWE 346
DB 269 FTWWTIARIHFEDYGLLRCDWNTINSSLWIIKGPILTSILVNFILFICIRILLQKLRP 328
QY 347 TNAVGHTRKQ-----YKRLAKSTLVLVFGVHYIVFVCLPHSTGLGWETRMHCELFN 402
DB 329 P-----DIRKSDSPYSRLARSTLLIPLFGVHYIMEAFFPDNFKP---EVKVMFELVVG 380
QY 403 SFQGFYIYCYCNGEVAQEVKMKWSRWNL--VDWK---RTPPCGSR---RCG---SVL 452
DB 381 SFQGFVVAIYCYCNGEVAQELRRKRRWHLQGLVGNPKYRHPSPGSGSGNGATCSTQVSM 440
QY 453 TTVT-----HSTSSQSOVA 466
DB 441 TRVSPGARRSSSQAEVS 458

RESULT 11

JC2195
vasoactive intestinal peptide receptor-related protein precursor (clone HIVR5) - human
C:Species: Homo sapiens (man)
C:date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 24-Nov-1999

C:Accession: JC2195; S42087
R:Biochem. Biophys. Res. Commun. 200, 769-776, 1994
A:Title: Human intestinal VIP receptor: cloning and functional expression of two cDNA
A:Reference number: JC2194; MUID:94235025
A:Accession: JC2195
A:Molecule type: mRNA
A:Residues: 1-495 <COU>
A:Cross-references: EMBL:X77777; NID:g456352; PIDN:CAA54814.1; PID:g456353
A:Experimental source: jejunal epithelial cell
C:Genetics:

A:Gene: GDB:VIPRI; RCD1; HVR1
A:Cross-references: GDB:128589; OMIM:192321
A:Map position: 3p22-3p22
C:Superfamily: glucagon receptor
C:Keywords: glycoprotein; intestine; phosphoprotein; receptor; transmembrane protein
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-495/Product: vasoactive intestinal peptide receptor-related protein #status predicted
F:180-203/Domain: transmembrane #status predicted <TM1>
F:211-229/Domain: transmembrane #status predicted <TM2>
F:251-269/Domain: transmembrane #status predicted <TM3>
F:290-312/Domain: transmembrane #status predicted <TM4>
F:334-354/Domain: transmembrane #status predicted <TM5>
F:381-398/Domain: transmembrane #status predicted <TM6>
F:412-431/Domain: transmembrane #status predicted <TM7>
F:93,104,135,328/Binding site: carboxylate (Asn) (covalent) #status predicted
F:111/Binding site: phosphate (Thr) (covalent) (by protein kinase A) #status predicted
F:285/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted
F:485/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 26.0%; Score 755.5; DB 2; Length 495;
Best Local Similarity 36.6%; Pred. No. 7.1e-56;
Matches 177; Conservative 80; Mismatches 138; Indels 89; Gaps 19;

QY 22 ARAQLDSGDTITIEQIVLVKAKVQCELNI-TAQLEGEGNCFPEWDGLICHPRTGTVK 80
DB 60 ARSLIGS-----SLQEECDYVQMIQKQLEEAQLENETIGCSKMDNLTCPATPRGQ 115
QY 81 ISAVPCPPYIYDNH---KGVAFPHCNPTWDFMHSLNKTNWYNSDCLRFLOPD----- 132
DB 116 VVILACPLIKFLFSSTQGRNVS-RSCTDEG-----WTH-----LEPGPYPIA 156
QY 133 -----ISIGKOE--FCERLYVMYTVGYSISFGLSAVALIIGYFRLHCTRNYIHHML 183
DB 157 CGLDDKAASLDEQOTWFYGSVKTYTIGYGLSLATLLVATILSLPRKLHCTRNYIHHML 216
QY 184 FVSFMLRATSFIVKDRVVAHIGVKELESIMQDDPONSTEATPSVDKSOYIGCKIAVVMF 243
DB 217 FISFILRAAAVFIKDLALF-----DSGESDQCSBGS-VGCKAAWVF 257
QY 244 IYFLATNYWILVEGLYLHNLIFVAFPSDTKYLWGFTILGWGPAAFAVAWAVARATLAD 303
DB 258 QYCVMANFWLLVEGLYLYTLAVSFPSERKYPFWGYLILGWVPSTFTWWTIARIHFED 317
QY 304 ---ARCWELSGADIKTYQAPILAAIGLNFILFNTVRLVATKIWEHCHTRKQ--- 357
DB 318 YGLLRCDWNTINSSLWIIKGPILTSILVNFILFICIRILLQKLRP-----DIRKSDSS 372
QY 358 -YKRLAKSTLVLVFGVHYIVFVCLPHSTGLGWETRMHCELFNFSQGFYIYCYC 416
DB 373 PYSRLARSTLLIPLFGVHYIMEAFFPDNFKP---EVKVMFELVVGSGFQGVVAIYCYFL 429
QY 417 NGEVQAEVKKMKWSRWNL--VDWK---RTPPCGSR---RCG---SVLTVT---HSTSSQ 462
DB 430 NGEVQAEELRRKRRWHLQGLVGNPKYRHPSPGSGSGNGATCSTQVSMILTRVSPGARRSSSQ 489
QY 463 SOVA 466
DB 490 AEVS 493

RESULT 12

A:Cross-references: EMBL:U20178; NID:g662795; PIDN:AA50106.1; PID:g662796
A:Experimental source: pancreas
C:Genetics:
A:Gene: GDB:SCTR
A:Cross-references: GDB:270546; OMIM:182098
A:Map position: 2q14.1-q14.1
C:Superfamily: glucagon receptor

```
Query Match          26.7%; Score 776; DB 2; Length 440;
Best Local Similarity 38.1%; Pred. No. 1.2e-57;
Matches 177; Conservative 64; Mismatches 139; Indels 84; Gaps 14;

Qy 13 WMLGSL--LARAQLSDGTTIEQIVLVLRKAVOCELNITAOEGEGNCFFPNDGL 70
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39 WEEDQCLQELSREQTGLDT---EQPV-----PCGEGMNDNI 73

Qy 71 ICWPRGTGKISAVCPPIYDF--NHKGVAFRHCNPNGTDFMHSLNKTNWYSDCLRF 129
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74 SCWSSVPGRMVEVECPRLRLTNRGSLFRNCTQDG-----WSE-----TFP 117

Qy 130 QPDISIG-----KQECERLYVMYTVGYSIFSGSLAVAILIIGYFRLHCTRNYI 179
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 RPNLACAVNVNDSEKHSLLKLVMTVGYSSSLVMLLVALLGILCAFRHLHCTRNYI 177

Qy 180 HMLHFSVPLRATSIYFKDRVVHAHIGVKELESIMQDDPONSIEATSVDKSQVIGCKIA 239
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 178 HMLHFSVPLRALSNFIKDAV-----LFSSDD-----VYYCD-AHRAGCKLV 218

Qy 240 VMETIYFLATNYWILVEGLYHLNLIYVAFSDTKYLMGFTLIGWGPFAAFVAWAVARA 299
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 219 MVLFQYCTIMANYSWLLVEGLYHLTLAISFSEKYLQGFVAFGWSIPAIFVALWAIARH 278

Qy 300 TLADARCWLSA-GDIKIYQAPILAAIGLNFILFLNTVRVLATKIWETNAVGHDTKQY 358
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 279 FLEDVGCMDINANASIMWIIRGPVLSLINFILFILMKRLKRTQETRGNEV-SHY 337

Qy 359 RKLAKSTLVLVVFGVHYIVVCLPHSFTGLGWIRHMCLEFFNSFGFFYSIIYCYNG 418
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 338 KRLARSTLLPLPGIHVYVAFSPED---AMEIQLFELALASFOGLVAVLYCFLNG 393

Qy 419 EVQAEVKMWSRWNLVDVKKTPPCGSRRCGSVLTVTHTSTSQ 462
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 394 EVQLVEQKKWQWHL-----REFPLHPVASFNSNKAHLEQSQ 432
```

RESULT 9
JH0594
vasoactive intestinal peptide receptor precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999
C:Accession: JH0594; S56014
R:Ishihara, T.; Shigemoto, R.; Mori, K.; Takahashi, K.; Nagata, S.
Neuron 8, 811-819, 1992
A:Title: Functional expression and tissue distribution of a novel receptor for vasoactive intestinal peptide
A:Reference number: JH0594; MUID:92232309
A:Accession: JH0594
A:Molecule type: mRNA
A:Residues: 1-459 <ISH>
A:Cross-references: GB:M86835; NID:g207640; PIDN:AAA42331.1; PID:g207641
A:Experimental source: lung
R:Pei, L.; Melmed, S.
Biochem. J. 308, 719-723, 1995
A:Title: Characterization of the rat vasoactive intestinal polypeptide receptor gene 5'
A:Reference number: S56014; MUID:97104266
A:Accession: S56014
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-26 <PEI>
A:Cross-references: EMBL:U10635; NID:g505752; PIDN:AAB48185.1; PID:g514311
C:Superfamily: glucagon receptor
C:Keywords: G protein-coupled receptor; glycoprotein; intestine; transmembrane protein
F:1-30/Domain: signal sequence #status predicted <SIG>

F:31-459/Product: vasoactive intestinal polypeptide receptor #status predicted <MAT>
F:146-168/Domain: transmembrane #status predicted <TM1>
F:176-195/Domain: transmembrane #status predicted <TM2>
F:218-241/Domain: transmembrane #status predicted <TM3>
F:256-277/Domain: transmembrane #status predicted <TM4>
F:295-318/Domain: transmembrane #status predicted <TM5>
F:344-363/Domain: transmembrane #status predicted <TM6>
F:376-395/Domain: transmembrane #status predicted <TM7>
F:58,69,100,292/Binding site: carbohydrate (Asn) (covalent) #status predicted

```
Query Match          26.6%; Score 772.5; DB 2; Length 459;
Best Local Similarity 36.8%; Pred. No. 2.4e-57;
Matches 172; Conservative 78; Mismatches 134; Indels 83; Gaps 17;

Qy 36 EQIVLVLRKAKVQCELNITAOEGEGNCFFPMDGLICWPHCTVCKISAVPCP-----PY 89
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38 EYLOLIEIQOCCLEE--AQLENETTGCSKMDNLCTPTTPRGOAVVLDCLPLIFQLFAP 95

Qy 90 IYDENHKGVAFRHCNPNGTDFMHSLNKTNWYSDCLRFLOP---DISIG----- 136
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 96 IHGYN----ISRSCTEG-----NSQ-----LEPGPYHIACGLNDRASSLDE 133

Qy 137 --KQECERLYVMYTVGYSIFSGSLAVAILIIGYFRLHCTRNYIHMLHFSVPLRATSI 194
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 134 QOQTKFYNTVGTGYTIGYSLASLLVAMATLSLFRKLHCTRNYIHMLHFSFILRATAV 193

Qy 195 FVKDRVVHAHIGVKELESIMQDDPONSIEATSVDKSQYIGCKTAVVMFYFLATNYWI 254
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 194 FIKDMALF-----NSGETDHCSEAS-VGCKAAVVFQYCVMANEFWL 234

Qy 255 LVEGLYLHNLIFVAFSDTKYLMGFTLIGWGPFAAFVAWAVARATLADARCWE-LSAGD 313
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 LVEGLYLTLTLLAVSFFSERKYEFGYLLIGWGPVSFVTITWTVRIYDFDCGCTIINSS 294

Qy 314 IKWTYQAPILAAIGLNFILFLNTVRVLATKIWETNAVGHDTKQYRKLAKSTLVLVVFG 373
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 LMWIKAPILLSLVNFVLCFICIRILVQKLRRPD-IGKNDSSPSYSLAKSTLTLPLFG 353

Qy 374 VHYIVFVCLPHSFTGLGWIRHMCLEFFNSFGFFYSIIYCYNGEQAQVKKMWSRWNL 433
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 354 IHYVMFAFFPDNFKA---QVMKVFELVVGSGFVFAILYFLNGEVAQLRRKWRMHL 410

Qy 434 S--VDW--KRTPPCGSRR---CG---SVLTVT---HSTSSQSOVA 466
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 411 QCVLWSSKSHQHPWGGSGNGATCTQVSMILFRVSPSARRSSSFQAEVS 457
```

RESULT 10
JC2194
vasoactive intestinal peptide receptor precursor - human
C:Species: Homo sapiens (man)
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 05-Nov-1999
C:Accession: JC2194; JN0604; PC2289; S38397
R:Couvineau, A.; Rouyer-Fessard, C.; Darmon, D.; Maoret, J.J.; Carro, I.; Ogier-De
Biochem. Biophys. Res. Commun. 200, 769-776, 1994
A:Title: Human intestinal VIP receptor: cloning and functional expression of two cDNA
A:Reference number: JC2194; MUID:94235025
A:Accession: JC2194
A:Molecule type: mRNA
A:Residues: 1-460 <COU>
A:Cross-references: EMBL:X75299; NID:g407461; PIDN:CAA53046.1; PID:g407462
A:Experimental source: jejunal epithelial cell; clone HVR8
R:Sreedharan, S.P.; Patel, D.R.; Huang, J.X.; Goetzl, E.J.
Biochem. Biophys. Res. Commun. 193, 546-553, 1993
A:Title: Cloning and functional expression of a human neuroendocrine vasoactive intes
A:Reference number: JN0604; MUID:93290641
A:Accession: JN0604
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-284,288-460 <SRE>
A:Cross-references: GB:L13288; NID:g292903; PIDN:AAA36805.1; PID:g292904
R:Couvineau, A.; Gaudin, P.; Maoret, J.J.; Rouyer-Fessard, C.; Nicole, P.; Laburthe,

QY	157	GSLAVAILIIGYFRRLLHCTRNYYTHMLFVSPFMLRATSIYFKORVVHAHHGVKELBSL----	213
		: : : :	
Db	200	ASLTAVAILIAYFPRLHCTRNYYTHMLFLSPMLRAVSIFVKDAVLVSGLTDLAEARLTEE	259
		: : : :	
QY	214	----IMODDPONSIEATSVDKSOYIGCKTAVNVIYFLATNYTWIIIVGLLYLHNLIFFAF	269
		: : : : : : : : : : : : :	
Db	260	ELRAIAQAPPPATAAG-----YACGRVAVTFEFLYLATNYTWIIIVGLYLHLSIFFAF	314
		: : : : : : : : : : : : :	
QY	270	FSDTKYLWGFIILGCGFPFAAFAAMAVARATLADARCWELSAGDIKWIQAPILAAIGIN	329
		: : : : : : : : : : : : :	
Db	315	FSEKKYLWGFTVGWLPAVFVAVMVSVRATLANTGCWDLSGNKKWIIQVPILASIVLN	374
		: : : : : : : : : : : : :	
QY	330	FILPLNTVRVLAKIWNETHNAVGHDRKRQYRKLAKTSLVLVLVFGVHYIVFVCLPHS-FTG	388
		: : : : : : : : : : : : :	
Db	375	FILFINIVRVLATKLRETNAGRCDTRQQYRKLLKSTLVLMPLFGVHYIVFMATPYTEVSG	434
		: : : : : : : : : : : : :	
QY	389	LGWIRMICLFPSNGFGFVSTIICYCNGEVOAEVKKMSRWNLISVDWKTPPCSRRC	448
		: : : : : : : : : : : : :	
Db	435	TLMQVQHMYEMLENSFGGFVALLIYCFNGEVOAEIKKSWSRTVALDFKKRARGSSSY	494
		: : : : : : : : : : : : :	
QY	449	G-SVLTVTHTSTSQSQAHAAMWCLSLAKLPSPADSITATSL-YLAMSGVTQTQR--AS	505
		: :	
Db	495	SYGPMSVSHTSVTVNGPRVGLG-----LPLSPRLPTATTNGHPQLPGHAKPGTAL	545
		: : : : : : : : : : : : :	
QY	506	HTLSTRS-----NKEDSGRQRDDIIMEK	528
		: : : : : : : : : : : : :	
Db	546	ETLETTPPMAAPKDDGFLNGSCGLUDEABEGPERPALLOEQ	587
		: : : : : : : : : : : : :	
RESULT	4		
I54195			
parathyroid hormone/parathyroid hormone related-peptide receptor - rat			
C:Species: Rattus norvegicus (Norway rat)			
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 11-Jan-2000			
C:Accession: I54195; A42698			
R:Pausova, Z.; Bourdon, J.; Clayton, D.; Mattei, M.G.; Seldin, M.F.; Janicic			
Genomics 20, 20-26, 1994			
A>Title: Cloning of a parathyroid hormone/parathyroid hormone-related peptide			
and rat genomes.			
A:Reference number: I54195; MUID:94292182			
A:Accession: I54195			
A>Status: preliminary; translated from GB/EMBL/DDBJ			
A:Molecule type: mRNA			
A:Residues: 1-591 <RES>			
A:Cross-references: GB:I19475; NID:g467316; PTDN:AAA68098.1; PID:g467317			
R:Abou-Samra, A.B.; Guppner, H.; Force, T.; Freeman, M.W.; Kong, X.F.; Schil			
proc. Natl. Acad. Sci. U.S.A. 89, 2732-2736, 1992			
A>Title: Expression cloning of a common receptor for parathyroid hormone and			
n of both cAMP and inositol trisphosphates and increases intracellular free			
A:Reference number: A42698; MUID:92212903			
A:Accession: A42698			
A>Status: preliminary; not compared with conceptual translation			
A:Molecule type: DNA			
A:Residues: 1-585, G', 587-591 <ARO>			
A:Experimental source: ROS 17/2.8 osteosarcoma cells			
A>Note: sequence extracted from NCBI backbone (NCBIP:92187)			
C:Superfamily: glucagon receptor			

Query Match	45.8%	Score 1331.5	DB 2	Length 591
Best Local Similarity	46.6%	Pred. No. 2.8e-104		
Matches 270	Conservative 86	Mismatches 136	Indels 87	Gaps
QY	22	ARQLSDSGTITTEQIIVLVKAKVOCE-----LNIITAQLORGEGN-----	62	
		: : : : : : : :		
Db	22	AYALVDADDVFTKEEQIFLLHRAQAQCCKLLKEVLHTAANIMESDKGWTGPASTSGPKRKE	81	
QY	63	-----CFPEWGLICWPRTGKLSAVPCPPIYIDFNHK	96	
Db	82	KASGKYPESKEKNQDVPYTSRRGRCLPSEWNIWCVLGAPEVVAVPCPPIYIDFNHK	141	
QY	97	GVAFRHCNPNGTWDFMHSLNKTTWANYSDCLRFLOPDISIGKGFCEGRLYVMYTVTGVSISF	156	

[illegible]

Query Match	45.5%	Score 1324;	DB 2;	Length 591;
Best Local Similarity	52.4%;	Pred. No. 1.2e-103;		
Matches 250;	Conservative	70;	Mismatches 97;	Indels 60;
			Gaps	6;

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:47:29 ; Search time 99.87 Seconds
(without alignments)
343.767 Million cell updates/sec

Title: US-09-236-468a-2
Perfect score: 2907
Sequence: 1 MAWLGASLHVWGMLGSL.....DDILMEKFSRPMESNPDTEG 541

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 182106

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2635	90.6	550	A57519	parathyroid hormon
2	1392	47.9	585	A39286	parathyroid hormon
3	1336.5	46.0	593	A49191	parathyroid hormon
4	1331.5	45.8	591	I54195	parathyroid hormon
5	1324	45.5	591	S44203	parathyroid hormon
6	1298.5	44.7	589	I59297	parathyroid hormon
7	801	27.6	449	S16319	secretin receptor
8	776	26.7	440	JC2532	secretin receptor
9	772.5	26.6	459	JH0594	vasoactive intesti
10	756	26.0	460	JC2194	vasoactive intesti
11	755.5	26.0	495	JC2195	vasoactive intesti
12	713.5	24.5	437	JU0185	PACAP/VP receptor
13	708	24.4	437	S39069	vasoactive intesti
14	702	24.1	438	G02822	vasoactive intesti
15	696.5	24.0	467	JN0616	pituitary adenylat
16	689.5	23.7	525	JN0902	pituitary adenylat
17	682.5	23.5	495	S39061	pituitary adenylat
18	682	23.5	494	S39063	pituitary adenylat
19	681.5	23.4	495	S36114	pituitary adenylat
20	675.5	23.2	495	A48204	pituitary adenylat
21	673.5	23.0	513	S47631	pituitary adenylat
22	667.5	23.0	523	S39060	pituitary adenylat
23	662	22.8	462	JC2462	gastric inhibitory
24	655	22.5	463	A46172	glucagon-like pept
25	641.5	22.1	466	G02234	gastric inhibitory
26	641.5	22.1	466	S66676	glucose-dependent
27	630.5	21.7	463	S18494	glucagon-like pept
28	628.5	21.6	423	A45363	somatoliberein rece
29	622.5	21.4	463	JN0807	glucagon-like pept

30	620.5	21.3	463	2	S71624	glucagon-like pept
31	619	21.3	477	2	JC2041	glucagon receptor
32	618.5	21.3	381	2	S33449	pituitary adenylat
33	616	21.2	455	2	I53273	gastric inhibitory
34	612	21.1	451	2	I46586	growth hormone-rel
35	607	20.9	485	2	JQ1957	glucagon receptor
36	604	20.8	491	2	I37411	glucose-dependent
37	596.5	20.5	423	2	S29753	growth hormone-rel
38	595.5	20.5	474	2	I37217	calcitonin recepto
39	595	20.5	485	2	JC4363	glucagon receptor
40	589.5	20.3	515	2	I49154	calcitonin recepto
41	583.5	20.1	479	2	S33746	calcitonin recepto
42	579.5	19.9	490	2	S34486	calcitonin recepto
43	579	19.9	478	2	A37430	calcitonin recepto
44	576.5	19.8	515	2	I60800	calcitonin recepto
45	575	19.8	464	2	S29754	growth hormone-rel

ALIGNMENTS

RESULT 1

A57519

parathyroid hormone receptor 2 precursor - human

N:Alternate names: PTH2 receptor

C:Species: Homo sapiens (man)

C>Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 24-Nov-1999

C:Accession: A57519

R:Usdin, T.B.; Gruber, C.; Bonner, T.I.

J. Biol. Chem. 270, 15455-15458, 1995

A:Title: Identification and functional expression of a receptor selectively recognizi

A:Reference number: A57519; MUID:95318121

A:Accession: A57519

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-550 <USD>

A:Cross-references: GB:U25128; NID:g887966; PIDN:AAC50157.1; PID:g887967

C:Genetics:

A:Gene: GDB: PTHR2; PTHR2R

A:Cross-references: GDB:731977; OMIM:601469

A:Map position: 2q33-2q33

C:Superfamily: glucagon receptor

C:Keywords: hormone receptor

Query Match 90.6%; Score 2635; DB 2; Length 550;

Best Local Similarity 91.4%; Pred. No. 4.6e-214;

Matches 500; Conservative 6; Mismatches 29; Indels 12; Gaps 3;

Qy 1 MAWLGASLHVWGMLGSLARALQSDGPTITIEEQIVLVKAKVOCELNITLQLOEGE 60

Db 1 MAGLCSLHVWGMLGSLARALQSDGPTITIEEQIVLVKAKVOCELNITLQLOEGE 60

Qy 61 GNCPEWDGLICPRGVTKISAVPCPPYIYDFNHKGVAFRHCNPNGTWDFMHSINKTWA 120

Db 61 GNCPEWDGLICPRGVTKISAVPCPPYIYDFNHKGVAFRHCNPNGTWDFMHSINKTWA 120

Qy 121 NYSCLRLPQDISIGKQECPCERLYVMYTVGYSTSFGLAVAILIIGYFRRLHCTRNVIH 180

Db 121 NYSCLRLPQDISIGKQECPCERLYVMYTVGYSTSFGLAVAILIIGYFRRLHCTRNVIH 180

Qy 181 MHLFVSFMLRATSFVKDRVVHAHGVKLESLIMODDPQNSIEATSVDKSQYIGCKTAV 240

Db 181 MHLFVSFMLRATSFVKDRVVHAHGVKLESLIMODDPQNSIEATSVDKSQYIGCKTAV 240

Qy 241 VMFYFLATNYIWLVEGLYLHNLIFFVAFFSDTKYLMGFIILIGWGFPAFVAANAVARAT 300

Db 241 VMFYFLATNYIWLVEGLYLHNLIFFVAFFSDTKYLMGFIILIGWGFPAFVAANAVARAT 300

Qy 301 LADARCWELSGADTKWIYQAPILAAIGLNFTLVTVRVLATKIWTETNAVGHDRKQYRK 360

Db 301 LADARCWELSGADTKWIYQAPILAAIGLNFTLVTVRVLATKIWTETNAVGHDRKQYRK 360

THIS PAGE BLANK (USPTO)

Db 347 P---IGISSTYOILFELCVGSGFQGLWAVLYCFLNSEVQCCLKRRWR----- 390
Qy 443 CGSRRCGSLVJTTHSTSSQSOVAAAHAWCLSLAKLPRSPADSLTATSLYLAMSGVTQSR 502
Db 391 -----GLCL-----TQGRSDYRLHSWMS-----RNGSES-----ALQIHGSRTOF 429
Qy 503 TASH 507
Db 430 LQSET 434

Search completed: November 8, 2000, 08:49:14
Job time: 107 sec

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/437,466
; FILING DATE: 09-MAY-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-136 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: RPACAP-3/RVIP-2
; HYPOTHETICAL: NO
; US-08-538-816A-2

Query Match          24.5%; Score 713.5; DB 2; Length 437;
Best Local Similarity 33.6%; Pred. No. 5.8e-66;
Matches 163; Conservative 87; Mismatches 156; Indels 79; Gaps 15;

QY 39 VLVLKAKVOCELNITAOQEGEGNCFP-----EWDGLICWPRGTGKISAV 84
Db 13 LLVRVSSIHPECRFHLIEQEEETKCAELLSSQENQACSGVWDNITCWRPADVGETTV 72

QY 85 PCPPYIYDF-NHKGVAFRHCNPNGTDFMHSKNTWANYSDCLRFLOPDISIGKQECER 143
Db 73 PCPKVFSNFYSRPGNISKNCTSDG-W-----SETPDFIDACGYNDPE-DESKISFYIL 124

QY 144 LYVMYTVGYSISFGSLAVAILIIGYFRRLHCTRNYIHMLFVSFMLRATSFYKDRVVHA 203
Db 125 VKAIYTLGYSVLSLTTGSIILCLFRKLHCTRNYIHLNLFSLFMLRAISVLVKDSVLYS 184

QY 204 HIGVKELESIMODDPQNSIEATSVDKSOYIGCKIAVVMFYIATNYIWLVEGLYLHN 263
Db 185 SSGI-----LRCHDQ-----ASWVGCKLSLVFFQYICIMANFYLLVEGLYLHT 228

QY 264 LIFVAFSDTKYLMGFTLIGWGFPAAFAVAVARATLADARCWELSGADIK-WIYQAPI 322
Db 229 LL-VAILPPSRCFYLALIGWGIPSCVIGAWTATRLSLEDTCGWDNDHSIPWVIRMPI 287

QY 323 LAAIGLNFILFNTVRVLATKIWETNAVGHDTKQYRKLAISTLVLVLFVGVHYIVFVCL 382
Db 288 LISIVNFALFISIVRLLQKLTSPDVGND-QSQYKRLAKSTLLLLPLFGVHYMVFAAF 346

QY 383 PHSFTGLGWEIRMHCELFNFSFGFFSIIYCYCNGEVAQEVKMKWSRWNLSDWKRTTP 442
Db 347 P-----IGISSTYQILFELCVGSGFGLVAVLYCFELNSEVQELKRRWR----- 390

QY 443 CGSRRCGSVLTTVHTSTSSQSOVAAAHACLSLAKLPRSPADSLTATSLYLAMSGVTQSR 502
Db 391 -----GLCL-----TQACSRDYLHWSMS-----RNGSES-----ALQIHRGSRQTQSF 429

QY 503 TASHIT 507
Db 430 LQSET 434

RESULT 15
US-09-076-651-2
; Sequence 2, Application US/09076651
; Patent No. 5882899
; GENERAL INFORMATION:
; APPLICANT: Mojssov, Svetlana
; APPLICANT: Wei, Yang
; TITLE OF INVENTION: RECEPTOR FOR PEPTIDE HORMONES INVOLVED
;
; TITLE OF INVENTION: IN ENERGY HOMEOSTASIS AND METHOD AND COMPOSITIONS FOR USE
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/076,651
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/538,816
; FILING DATE: 03-OCTOBER-1995
; APPLICATION NUMBER: US 08/437,466
; FILING DATE: 09-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-136 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: RPACAP-3/RVIP-2
; HYPOTHETICAL: NO
; US-09-076-651-2

Query Match          24.5%; Score 713.5; DB 2; Length 437;
Best Local Similarity 33.6%; Pred. No. 5.8e-66;
Matches 163; Conservative 87; Mismatches 156; Indels 79; Gaps 15;

QY 39 VLVLKAKVOCELNITAOQEGEGNCFP-----EWDGLICWPRGTGKISAV 84
Db 13 LLVRVSSIHPECRFHLIEQEEETKCAELLSSQENQACSGVWDNITCWRPADVGETTV 72

QY 85 PCPPYIYDF-NHKGVAFRHCNPNGTDFMHSKNTWANYSDCLRFLOPDISIGKQECER 143
Db 73 PCPKVFSNFYSRPGNISKNCTSDG-W-----SETPDFIDACGYNDPE-DESKISFYIL 124

QY 144 LYVMYTVGYSISFGSLAVAILIIGYFRRLHCTRNYIHMLFVSFMLRATSFYKDRVVHA 203
Db 125 VKAIYTLGYSVLSLTTGSIILCLFRKLHCTRNYIHLNLFSLFMLRAISVLVKDSVLYS 184

QY 204 HIGVKELESIMODDPQNSIEATSVDKSOYIGCKIAVVMFYIATNYIWLVEGLYLHN 263
Db 185 SSGI-----LRCHDQ-----ASWVGCKLSLVFFQYICIMANFYLLVEGLYLHT 228

QY 264 LIFVAFSDTKYLMGFTLIGWGFPAAFAVAVARATLADARCWELSGADIK-WIYQAPI 322
Db 229 LL-VAILPPSRCFYLALIGWGIPSCVIGAWTATRLSLEDTCGWDNDHSIPWVIRMPI 287

QY 323 LAAIGLNFILFNTVRVLATKIWETNAVGHDTKQYRKLAISTLVLVLFVGVHYIVFVCL 382
Db 288 LISIVNFALFISIVRLLQKLTSPDVGND-QSQYKRLAKSTLLLLPLFGVHYMVFAAF 346

QY 383 PHSFTGLGWEIRMHCELFNFSFGFFSIIYCYCNGEVAQEVKMKWSRWNLSDWKRTTP 442
Db 347 P-----IGISSTYQILFELCVGSGFGLVAVLYCFELNSEVQELKRRWR----- 390

QY 443 CGSRRCGSVLTTVHTSTSSQSOVAAAHACLSLAKLPRSPADSLTATSLYLAMSGVTQSR 502
Db 391 -----GLCL-----TQACSRDYLHWSMS-----RNGSES-----ALQIHRGSRQTQSF 429

QY 503 TASHIT 507
Db 430 LQSET 434

RESULT 15
US-09-076-651-2
; Sequence 2, Application US/09076651
; Patent No. 5882899
; GENERAL INFORMATION:
; APPLICANT: Mojssov, Svetlana
; APPLICANT: Wei, Yang
; TITLE OF INVENTION: RECEPTOR FOR PEPTIDE HORMONES INVOLVED
;
; TITLE OF INVENTION: IN ENERGY HOMEOSTASIS AND METHOD AND COMPOSITIONS FOR USE
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/076,651
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/538,816
; FILING DATE: 03-OCTOBER-1995
; APPLICATION NUMBER: US 08/437,466
; FILING DATE: 09-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-136 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: RPACAP-3/RVIP-2
; HYPOTHETICAL: NO
; US-08-538-816A-2
```


; Sequence 5, Application US/08869477
; Patent No. 5846747
; GENERAL INFORMATION:
; APPLICANT: Thorens, Bernard
; TITLE OF INVENTION: Receptor for the Glucagon-Like-Peptide-1
; TITLE OF INVENTION: (GLP-1)
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5846747o No. 5846747disk of No. 5846747th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/869,477
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,439
; FILING DATE: 24-NOV-93
; APPLICATION NUMBER: DK 398/92
; FILING DATE: 25-MAR-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/00697
; FILING DATE: 23-MAR-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Harrington, James J.
; REGISTRATION NUMBER: 38,711
; REFERENCE/DOCKET NUMBER: 3756.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Rattus norvegicus
; STRAIN: Sprague-Dawley
; US-08-869-477-5

Query Match 27.6%; Score 801; DB 2; Length 449;
Best Local Similarity 37.3%; Pred. No. 4.6e-75;
Matches 176; Conservative 81; Mismatches 149; Indels 66; Gaps 15;
QY 14 LMLGSLARAOALDSGCTITIEQIVLVKAKVQCELNITAOQL-OEGEGNCFPE----- 66
Db 12 LLRLRLTLRAA-----HTVGVPRLCDVRVLLERAHCLQJLSKRRKKGALGPETASGCE 67
QY 67 --WDGLICWPRGTGKISAVPCPPYIYDFNHK-GVAFRHCPNPGTWDVFMHSLNKNWTNAYS 123
Db 68 GLMDNMSCWSPSSAPARTVEVQCPKFLMLSNKNGSLFRNCTQDG-----WSE-- 114
QY 124 DCLRFQPDISIG---KOEFCER-----LYWYTVGVYSISFGSLAVAILIIGYFRLH 173
Db 115 ---TFPRPDLCAGVININNSFNERRHAYLLKLVMTYTVGYSSSLAMLLVALSILCSFRLH 171
QY 174 CTRNYIIMHLFVSPMLRATSIYFKDRVVVHAHIGVKELESIMODDPONSIENTSVKSOY 233
Db 172 CTRNYIIMHLFVSPMLRATSIYFKDRVVVHAHIGVKELESIMODDPONSIENTSVKSOY 233
Db 172 CTRNYIIMHLFVSPMLRATSIYFKDRVVVHAHIGVKELESIMODDPONSIENTSVKSOY 233

QY 234 IGCKIAVVMFIYFLATNYWILVEGLYHLNLIFFVAFESDTKYLWGFILIGWFPFAFVAA 293
Db 213 VGCKLVMIFFOYICIMANYAMLLVBLVHTLLAISFFSERKYLQAFVLLGWSIPAIFVAL 272
QY 294 WAVARATLADARCWELSA-GDIKIYQAPILAAIGLNFILFNTVRVLATKINETNAVGH 352
Db 273 WAITRHFLENTGCDINANASVMVIRGVPVILSILNFIFFINILRLIMRKLRTQETRGS 332
QY 353 DTRKQYKRLAKSTLVLVLVFGVHVIVCLPHSFTGLGWEIRHMCFLFFNSFQGFVYSII 412
Db 333 ET-NHYKRLAKSTLVLVLVFGVHVIVCLPHSFTGLGWEIRHMCFLFFNSFQGFVYSII 412
QY 413 YCYNGEVOAEVKKMWSRWNLSDVKETPPCGSRRCGSLVTTVTHSTSSOSQ 464
Db 388 YCFLNGEVOAEVKKMWSRWNLSDVKETPPCGSRRCGSLVTTVTHSTSSOSQ 464
RESULT 12
US-08-112-817C-2
; Sequence 2, Application US/08112817C
; Patent No. 5573928
; GENERAL INFORMATION:
; APPLICANT: Hsiung, Hansen M.
; APPLICANT: Smith, Dennis P.
; APPLICANT: Zhang, Xing-Yue
; TITLE OF INVENTION: PORCINE VASOACTIVE INTESTINAL PEPTIDE
; TITLE OF INVENTION: RECEPTOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh Iici compatible
; OPERATING SYSTEM: System 7
; SOFTWARE: Microsoft Word for Macintosh v.5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/112,817C
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy, Richard B.
; REGISTRATION NUMBER: 35,296
; REFERENCE/DOCKET NUMBER: X-9293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3589
; TELEFAX: 317-276-1294
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 458 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-112-817C-2

Query Match 27.0%; Score 785; DB 1; Length 458;
Best Local Similarity 37.0%; Pred. No. 2.2e-73;
Matches 182; Conservative 76; Mismatches 152; Indels 82; Gaps 18;
QY 12 GWMLGSLCIA---RAOLDSGCTITI---BEQIVLVKAKVQCELNITAOQLGEGNCF 64
Db 10 GWF---CVLAGVLACVLPVPGVSWAVGLQOECDYLOMIKVOHKQCLEAEQENETSGCS 65
QY 65 PENDGLICWPRGTGKISAVPCPPYIYDFN-HKGV-AFRHCPNPGTWDVFMHSLNKNWTNANY 122
Db 66 KMDNLTCPATPGVQVVLACPLIFKLFSTQGLVSNRNCITDEG-----WTP- 113
QY 123 SDCLRFQPDISIG---KOEFCER-----LYWYTVGVYSISFGSLAVAILIIGY 168


```

; 144 RCDNSGSEWLVPGNNRTWYSECVKFLTNETR--EREVDFRLGMIYTVGYSISGLSLTV 201
; 162 AILIGYFRRLHCTRNIIHMLFVSFMRATSIIVKDRVVAHIGVKELESIMQDDPON 221
; 202 AVILIGYFRRLHCTRNIIHMLFVSFMRATSIIVKDRVVAHIGVKELESIMQDDPON 260
; 222 SIATSDVKSOYIGCKIAVVMYIYFLATNYWILVEGLYHNLIFVAFSDTKYLMGFIL 281
; 261 FTEPPADKAGFVGCRAVTVFLYLTNYWILVEGLYHNLIFVAFSDTKYLMGFIL 320
; 282 IGWGFPAFVAANAVARATLADARWELSGADIKWYQAPILAAIGLNFILFNTVRVLA 341
; 321 FGWGLPAFVAANAVARATLADARWELSGADIKWYQAPILAAIGLNFILFNTVRVLA 380
; 342 TKIWTNAVGHDRKQYRKLAFTLVLVFGVHYIVFVCLPHS--FTGLGWEIRMHCELF 400
; 381 TKLRETNAGRCDRQQYRKLLKSTLVLMPLFGVHYIVFVCLPHS--FTGLGWEIRMHCELF 440
; 401 FNSFGFFVSIYCYCNGEYQAEVKMWSRWNLSDVMKRTTPCGSRRCGSLVLTVTH-- 457
; 441 FNSFGFFVSIYCYCNGEYQAEVKMWSRWNLSDVMKRTTPCGSRRCGSLVLTVTH-- 493
; 458 -STSSQSOVAAAHAWCLSLAKLPRSPADSLTATSLYLAHSGVTSQRTASHTLSTRNKED 516
; 494 VSHTSVTNVGPRGGLALSLS--PRLAPGAGASANGHQLPGYVKHGSISEN-SLPSSGPE 550
; 517 SGRQRDDILM-----EKPSRPMESNPDT 539
; 551 PGTKDDGYLNGSLYEPWVGEPPLLEERET 583
```

RESULT 5

```

US-08-142-551B-125
; Sequence 125, Application US/08142551B
; Patent No. 5814603
; GENERAL INFORMATION:
; APPLICANT: Oldenburg, Kevin R.
; TITLE OF INVENTION: COMPOUNDS WITH PTH ACTIVITY AND
; TITLE OF INVENTION: RECOMBINANT DNA VECTORS ENCODING SAME
; NUMBER OF SEQUENCES: 132
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: US
; ZIP: 22313
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,551B
; FILING DATE: 25-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/077,296
; FILING DATE: 14-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/898,219
; FILING DATE: 12-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/965,677
; FILING DATE: 22-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Swiss, Gerald F.
; REGISTRATION NUMBER: 30,113
; REFERENCE/DOCKET NUMBER: 000324-010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-7400
```

```

; TELEFAX: (415) 854-8275
; INFORMATION FOR SEQ ID NO: 125:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..585
; OTHER INFORMATION: /note= "PTH receptor"
; US-08-142-551B-125

Query Match 47.9%; Score 1392; DB 2; Length 585;
Best Local Similarity 47.8%; Pred. No. 1.8e-136;
Matches 274; Conservative 92; Mismatches 137; Indels 70; Caps 9;

QY 24 AQLSDSGTITIEQIVLVKAKVQCELNITAIQLQEGE----- 60
Db 24 ALVDADDVITKEEQIILLRNAQACEQRLAEVLRVPELAESAKDMWSRSKATKKERPAEK 83
QY 61 -----GNCPEWDGLICWPRGTVGKISAVPCPPYIYDFNHKGVAFR 101
Db 84 LYPQAEESREVSDRSRLQDGFCLPEMDNIVCWPAQVGVAVPCPDYFYDFNHKGGRAY 143
QY 102 HCNPCNTWDFHSLINKTWANYSOCLRFLOPDISIGKQECERLYVMYTVGYSISFGSLAV 161
Db 144 RCDNSGSEWLVPGNNRTWYSECVKFLTNETR--EREVDFRLGMIYTVGYSISGLSLTV 201
QY 162 AILIGYFRRLHCTRNIIHMLFVSFMRATSIIVKDRVVAHIGVKELESIMQDDPON 221
Db 202 AVILIGYFRRLHCTRNIIHMLFVSFMRATSIIVKDRVVAHIGVKELESIMQDDPON 260
QY 222 SIATSDVKSOYIGCKIAVVMYIYFLATNYWILVEGLYHNLIFVAFSDTKYLMGFIL 281
Db 261 FTEPPADKAGFVGCRAVTVFLYLTNYWILVEGLYHNLIFVAFSDTKYLMGFIL 320
QY 282 IGWGFPAFVAANAVARATLADARWELSGADIKWYQAPILAAIGLNFILFNTVRVLA 341
Db 321 FGWGLPAFVAANAVARATLADARWELSGADIKWYQAPILAAIGLNFILFNTVRVLA 380
QY 342 TKIWTNAVGHDRKQYRKLAFTLVLVFGVHYIVFVCLPHS--FTGLGWEIRMHCELF 400
Db 381 TKLRETNAGRCDRQQYRKLLKSTLVLMPLFGVHYIVFVCLPHS--FTGLGWEIRMHCELF 440
QY 401 FNSFGFFVSIYCYCNGEYQAEVKMWSRWNLSDVMKRTTPCGSRRCGSLVLTVTH-- 457
Db 441 FNSFGFFVSIYCYCNGEYQAEVKMWSRWNLSDVMKRTTPCGSRRCGSLVLTVTH-- 493
QY 458 -STSSQSOVAAAHAWCLSLAKLPRSPADSLTATSLYLAHSGVTSQRTASHTLSTRNKED 516
Db 494 VSHTSVTNVGPRGGLALSLS--PRLAPGAGASANGHQLPGYVKHGSISEN-SLPSSGPE 550
QY 517 SGRQRDDILM-----EKPSRPMESNPDT 539
Db 551 PGTKDDGYLNGSLYEPWVGEPPLLEERET 583
```

RESULT 6

```

US-08-468-249A-19
; Sequence 19, Application US/08468249A
; Patent No. 5886148
; GENERAL INFORMATION:
; APPLICANT: Segre et al., Gino V.
; TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR AND DNA
; TITLE OF INVENTION: ENCODING SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,439A
; FILING DATE: 24-NOV-93
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 398/92
; FILING DATE: 25-MAR-92
; PRIOR APPLICATION DATA: PCT/EP93/00697
; APPLICATION NUMBER: 23-MAR-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Harrington, James J.
; REGISTRATION NUMBER: 38,711
; REFERENCE/DOCKET NUMBER: 3756.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Didelphis virginiana
; US-08-142-439A-6

Query Match 48.1%; Score 1397; DB 1; Length 585;
Best Local Similarity 48.0%; Pred. No. 5.5e-137;
Matches 275; Conservative 94; Mismatches 134; Indels 70; Gaps 10;

QY 24 AQLDSGDTITIEEQIVLVKAKVQCELNIT-----53
Db 24 ALVDADDVITKEQIILLRNAQACEQRLKEVLVPPELAESAKDMWSRSKTKKPAEK 83

QY 54 --AQLQEG-----EGNCFPEWDLICWPGRGTGKISAVPCPPYIYDFNHKGVAFR 101
Db 84 LYSQAEESREVSDRSRLQDGFCLPEWNIWCPAGVPGKVAVPCPDYIYDFNHKGRAYR 143

QY 102 HCNPNGTWDFMHSINKTWANYSDCLRFQPDISIGKQFCERLYVMYTVGYSISFGSLAV 161
Db 144 RCDNNGSWELVPGNNRTWANYSECVKPLTNETR--EREVDFRLGMIYTVGYSISGLSTV 201

QY 162 AILLGVFERHCTNVIHMLFVDEMLRATISFVKDRVVHAHIGVKELESIMQDDPON 221
Db 202 AVLLGVFRHCTNVIHMLFVDEMLRATISFVKDRVVHAHIGVKELESIMQDDPON 260

QY 222 SIKATSDVKSQYIGCKTAVVMFIYVLTATNYWILVEGLYLNHLIFVAFPSDTKYLWGFIL 281
Db 261 FTEPPPADKAGFVGRVAVTVFVFLTNTYWIIVVEGLYLSLIFMAFFSEKKYLGFTL 320

QY 282 IGWGFPAFVAANAVARATLADARCWELSAGDKTWIQAPILAAIGLNFILFLNTVRVLA 341
Db 321 FGWGLPAFVAVVAVTVVTLANTLANTECWDLSSGNKKWIIQVPIAAIVVNFILFINIIRVLA 380

QY 342 TKIWTNAVGHDRKQYRKLAKTSLVLVFGVHYIVFVCLPHS--FTGLGWELRMHCELF 400
Db 381 TKURETAGRCDFRQYRKLLKSTVLVPLFGVHYIVFVMTATPYTEVSGILWQVMHYEML 440

QY 401 FNSQGFVSIICYCNGEVAQAEVKKMWSRWNLSDVMKTRTPPCGSRRCGSLVLTVTTH-- 457
Db 441 FNSQGFVSIICYCNGEVAQAEVKKMWSRWNLSDVMKTRTPPCGSRRCGSLVLTVTTH-- 493

QY 458 -STSSQSQAHAAMWLUSLAKLPRSPADSLTATSLYLAMSGVTVQSTRTASHTLSRKNED 516
Db 494 VSHTSVTNVGPRGGLALSLS--PRLAPGAGASANGHQLPGYVVKHGSISEN-SLPSSGPE 550

QY 517 SGRQRDDILM-----EKPSRPMESNPDT 539
Db 551 PGTKDDGYLNGSGLYEPMVGEQPPPLLEERET 583

RESULT 4
US-08-869-477-6
; Sequence 6, Application US/08869477
; Patent No. 5846747
; GENERAL INFORMATION:
; APPLICANT: Thorens, Bernard
; TITLE OF INVENTION: Receptor for the Glucagon-Like-Peptide-1
; TITLE OF INVENTION: (GLP-1)
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5846747o No. 5846747disk of No. 5846747th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/869,477
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,439
; FILING DATE: 24-NOV-93
; APPLICATION NUMBER: DK 398/92
; FILING DATE: 25-MAR-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/00697
; FILING DATE: 23-MAR-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Harrington, James J.
; REGISTRATION NUMBER: 38,711
; REFERENCE/DOCKET NUMBER: 3756.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Didelphis virginiana
; US-08-869-477-6

Query Match 48.1%; Score 1397; DB 2; Length 585;
Best Local Similarity 48.0%; Pred. No. 5.5e-137;
Matches 275; Conservative 94; Mismatches 134; Indels 70; Gaps 10;

QY 24 AQLDSGDTITIEEQIVLVKAKVQCELNIT-----53
Db 24 ALVDADDVITKEQIILLRNAQACEQRLKEVLVPPELAESAKDMWSRSKTKKPAEK 83

QY 54 --AQLQEG-----EGNCFPEWDLICWPGRGTGKISAVPCPPYIYDFNHKGVAFR 101
Db 84 LYSQAEESREVSDRSRLQDGFCLPEWNIWCPAGVPGKVAVPCPDYIYDFNHKGRAYR 143

QY 102 HCNPNGTWDFMHSINKTWANYSDCLRFQPDISIGKQFCERLYVMYTVGYSISFGSLAV 161
Db 144 RCDNNGSWELVPGNNRTWANYSECVKPLTNETR--EREVDFRLGMIYTVGYSISGLSTV 201

QY 162 AILLGVFERHCTNVIHMLFVDEMLRATISFVKDRVVHAHIGVKELESIMQDDPON 221
Db 202 AVLLGVFRHCTNVIHMLFVDEMLRATISFVKDRVVHAHIGVKELESIMQDDPON 260

QY 222 SIKATSDVKSQYIGCKTAVVMFIYVLTATNYWILVEGLYLNHLIFVAFPSDTKYLWGFIL 281
Db 261 FTEPPPADKAGFVGRVAVTVFVFLTNTYWIIVVEGLYLSLIFMAFFSEKKYLGFTL 320

QY 282 IGWGFPAFVAANAVARATLADARCWELSAGDKTWIQAPILAAIGLNFILFLNTVRVLA 341
Db 321 FGWGLPAFVAVVAVTVVTLANTLANTECWDLSSGNKKWIIQVPIAAIVVNFILFINIIRVLA 380

QY 342 TKIWTNAVGHDRKQYRKLAKTSLVLVFGVHYIVFVCLPHS--FTGLGWELRMHCELF 400
Db 381 TKURETAGRCDFRQYRKLLKSTVLVPLFGVHYIVFVMTATPYTEVSGILWQVMHYEML 440

QY 401 FNSQGFVSIICYCNGEVAQAEVKKMWSRWNLSDVMKTRTPPCGSRRCGSLVLTVTTH-- 457
Db 441 FNSQGFVSIICYCNGEVAQAEVKKMWSRWNLSDVMKTRTPPCGSRRCGSLVLTVTTH-- 493

QY 458 -STSSQSQAHAAMWLUSLAKLPRSPADSLTATSLYLAMSGVTVQSTRTASHTLSRKNED 516
Db 494 VSHTSVTNVGPRGGLALSLS--PRLAPGAGASANGHQLPGYVVKHGSISEN-SLPSSGPE 550
```

```

Db 1 MAMGLASLHWGMLGSCLLARAQLDSDGTITIEQIVLVKAKVOCELNITAQLOEGE 60
QY 61 GNCFPWDGLICPRGTGKISAVPCPPYIYDFNHHKGVAFRHCNPNCTWDFMHSLNKTWA 120
Db 61 GNCFPWDGLICPRGTGKISAVPCPPYIYDFNHHKGVAFRHCNPNCTWDFMHSLNKTWA 120
QY 121 NYSDCLRFLOPDISIGKQECERLYVMYTVGYSISFGSLAVAILIIGYFRRLHCTRNYIH 180
Db 121 NYSDCLRFLOPDISIGKQECERLYVMYTVGYSISFGSLAVAILIIGYFRRLHCTRNYIH 180
QY 181 MHLFVSFMLRATSIYFKDRVHVHAHIGVKELESIMODDPONSIEATSVDKSOYIGCKIAV 240
Db 181 MHLFVSFMLRATSIYFKDRVHVHAHIGVKELESIMODDPONSIEATSVDKSOYIGCKIAV 240
QY 241 VMFYFLATNYWILVEGLYHNLIFVAFESDRTKYLWGFILGWGFPFAFVAWAVARAT 300
Db 241 VMFYFLATNYWILVEGLYHNLIFVAFESDRTKYLWGFILGWGFPFAFVAWAVARAT 300
QY 301 LADARCHELSAGDIKWYQAPILAAIGLNFILFNTVVRVLATKIWETNAVGHDTRKQYRK 360
Db 301 LADARCHELSAGDIKWYQAPILAAIGLNFILFNTVVRVLATKIWETNAVGHDTRKQYRK 360
QY 361 LAKSTLVLVFVGVHYIVFVCLPHSFITGLGWEIRMHCELFNFSGQFFVSIICYNGEV 420
Db 361 LAKSTLVLVFVGVHYIVFVCLPHSFITGLGWEIRMHCELFNFSGQFFVSIICYNGEV 420
QY 421 QAEVKKMSRWNLSDVMKRTPPCGSRRCGSLVTTVTHSTSSQSOVAAAHAACLSLAKLPR 480
Db 421 QAEVKKMSRWNLSDVMKRTPPCGSRRCGSLVTTVTHSTSSQSOVAAAHAACLSLAKLPR 480
QY 481 SPADSLTATSLYLAMSGVTSQRTASHTPLSTRNKEDSGRQDDILMEKPSRPMESNPDE 540
Db 481 SPADSLTATSLYLAMSGVTSQRTASHTPLSTRNKEDSGRQDDILMEKPSRPMESNPDE 540
QY 541 G 541
Db 541 G 541

```

RESULT 2

```

PCT-US95-07085-2
; Sequence 2, Application PC/TUS9507085
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone Receptor
; TITLE OF INVENTION: HLTG74
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESS: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07085
; FILING DATE: 05-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-393
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700

```

```

; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-07085-2

Query Match 100.0%; Score 2907; DB 4; Length 541;
Best Local Similarity 100.0%; Pred. No. 2.3e-294; Indels 0; Gaps 0;
Matches 541; Conservative 0; Mismatches 0;

QY 1 MAMGLASLHWGMLGSCLLARAQLDSDGTITIEQIVLVKAKVOCELNITAQLOEGE 60
Db 1 MAMGLASLHWGMLGSCLLARAQLDSDGTITIEQIVLVKAKVOCELNITAQLOEGE 60
QY 61 GNCFPWDGLICPRGTGKISAVPCPPYIYDFNHHKGVAFRHCNPNCTWDFMHSLNKTWA 120
Db 61 GNCFPWDGLICPRGTGKISAVPCPPYIYDFNHHKGVAFRHCNPNCTWDFMHSLNKTWA 120
QY 121 NYSDCLRFLOPDISIGKQECERLYVMYTVGYSISFGSLAVAILIIGYFRRLHCTRNYIH 180
Db 121 NYSDCLRFLOPDISIGKQECERLYVMYTVGYSISFGSLAVAILIIGYFRRLHCTRNYIH 180
QY 181 MHLFVSFMLRATSIYFKDRVHVHAHIGVKELESIMODDPONSIEATSVDKSOYIGCKIAV 240
Db 181 MHLFVSFMLRATSIYFKDRVHVHAHIGVKELESIMODDPONSIEATSVDKSOYIGCKIAV 240
QY 241 VMFYFLATNYWILVEGLYHNLIFVAFESDRTKYLWGFILGWGFPFAFVAWAVARAT 300
Db 241 VMFYFLATNYWILVEGLYHNLIFVAFESDRTKYLWGFILGWGFPFAFVAWAVARAT 300
QY 301 LADARCHELSAGDIKWYQAPILAAIGLNFILFNTVVRVLATKIWETNAVGHDTRKQYRK 360
Db 301 LADARCHELSAGDIKWYQAPILAAIGLNFILFNTVVRVLATKIWETNAVGHDTRKQYRK 360
QY 361 LAKSTLVLVFVGVHYIVFVCLPHSFITGLGWEIRMHCELFNFSGQFFVSIICYNGEV 420
Db 361 LAKSTLVLVFVGVHYIVFVCLPHSFITGLGWEIRMHCELFNFSGQFFVSIICYNGEV 420
QY 421 QAEVKKMSRWNLSDVMKRTPPCGSRRCGSLVTTVTHSTSSQSOVAAAHAACLSLAKLPR 480
Db 421 QAEVKKMSRWNLSDVMKRTPPCGSRRCGSLVTTVTHSTSSQSOVAAAHAACLSLAKLPR 480
QY 481 SPADSLTATSLYLAMSGVTSQRTASHTPLSTRNKEDSGRQDDILMEKPSRPMESNPDE 540
Db 481 SPADSLTATSLYLAMSGVTSQRTASHTPLSTRNKEDSGRQDDILMEKPSRPMESNPDE 540
QY 541 G 541
Db 541 G 541

RESULT 3
US-08-142-439A-6
; Sequence 6, Application US/08142439A
; Patent No. 5670360
; GENERAL INFORMATION:
; APPLICANT: Thorens, Bernard
; TITLE OF INVENTION: Receptor for the Glucagon-Like-Peptide-1
; TITLE OF INVENTION: (GLP-1)
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5670360 No. 5670360disk of No. 5670360th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:47:27 ; Search time 97.15 Seconds
(without alignments)
93.338 Million cell updates/sec

Title: US-09-236-468A-2
Perfect score: 2907
Sequence: 1 MAWLGCASLHVWGLMGSL.....DDLMEKPSRPMSNPDEG 541

Scoring table: BLOSUM62
Gapop 10.0 , Grpext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 164575

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2907	100.0	541	3	US-08-468-011A-2
2	2907	100.0	541	4	PCT-US95-07085-2
3	1397	48.1	585	1	US-08-142-439A-6
4	1397	48.1	585	2	US-08-869-477-6
5	1392	47.9	585	2	US-08-142-551B-125
6	1392	47.9	585	2	US-08-468-249A-19
7	1388	47.7	535	2	US-08-468-249A-18
8	1331.5	45.8	591	2	US-08-468-249A-20
9	1331.5	45.8	593	2	US-08-468-249A-21
10	801	27.6	449	1	US-08-142-439A-5
11	801	27.6	449	2	US-08-869-477-5
12	785	27.0	458	1	US-08-112-817C-2
13	773	26.6	1324	2	US-08-811-897A-56
14	713.5	24.5	437	2	US-08-538-816A-2
15	713.5	24.5	437	2	US-09-076-651-2
16	704	24.2	431	2	US-08-538-816A-9
17	704	24.2	431	2	US-09-076-651-9
18	702.5	24.2	438	2	US-08-538-816A-1
19	702.5	24.2	438	2	US-09-076-651-1
20	696.5	24.0	448	2	US-08-811-897A-18
21	696.5	24.0	448	2	US-08-855-213-18
22	696.5	24.0	467	2	US-08-811-897A-19
23	696.5	24.0	467	2	US-08-855-213-19
24	689.5	23.7	525	2	US-08-811-897A-23
25	689.5	23.7	525	2	US-08-855-213-23
26	688.5	23.7	448	2	US-08-811-897A-16
27	688.5	23.7	448	2	US-08-855-213-16
28	688.5	23.7	485	2	US-08-811-897A-17

29	688.5	23.7	485	2	US-08-855-213-17	Sequence 17, Appl
30	684.5	23.5	448	2	US-08-811-897A-22	Sequence 22, Appl
31	684.5	23.5	448	2	US-08-855-213-22	Sequence 22, Appl
32	681.5	23.4	476	2	US-08-811-897A-20	Sequence 20, Appl
33	681.5	23.4	476	2	US-08-855-213-20	Sequence 20, Appl
34	681.5	23.4	495	2	US-08-811-897A-21	Sequence 21, Appl
35	681.5	23.4	495	2	US-08-855-213-21	Sequence 21, Appl
36	675	23.2	552	2	US-08-811-897A-27	Sequence 27, Appl
37	675	23.2	552	2	US-08-855-213-27	Sequence 27, Appl
38	674.5	23.2	553	2	US-08-811-897A-25	Sequence 25, Appl
39	674.5	23.2	553	2	US-08-855-213-25	Sequence 25, Appl
40	673.5	23.2	476	2	US-08-811-897A-14	Sequence 14, Appl
41	673.5	23.2	476	2	US-08-855-213-14	Sequence 14, Appl
42	673.5	23.2	513	2	US-08-811-897A-15	Sequence 15, Appl
43	673.5	23.2	513	2	US-08-855-213-15	Sequence 15, Appl
44	672.5	23.1	553	2	US-08-811-897A-29	Sequence 29, Appl
45	672.5	23.1	553	2	US-08-855-213-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-08-468-011A-2
; Sequence 2, Application US/08468011A
; Patent No. 6030804
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R
; APPLICANT: Yi, Li
; APPLICANT: Rosen, Craig A
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: G-Protein Parathyroid Hormone receptor
; TITLE OF INVENTION: ILTDG74
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,011A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-458 (PF201)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 541 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-011A-2

Query Match 100.0%; Score 2907; DB 3; Length 541;
Best Local Similarity 100.0%; Pred. No. 2.3e-294;
Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAWLGCASLHVWGLMGSLARAQSDGCTITIEEQIVLVKAKVOCENITAAQLOEGE 60
|||||

THIS PAGE BLANK (USPTO)

Query Match 27.6%; Score 801; DB 13; Length 449;
Best Local Similarity 37.3%; Pred. No. 2.6e-78;
Matches 176; Conservative 81; Mismatches 149; Indels 66; Gaps 15;

QY 14 LMLGSCILARQLDSGDTITTEQIVLVKAKVQCELNITAOQ-OEGRCNCFPE----- 66
DB 12 lllrlllllkkaa---hvgvpprldcrrvllleerachclgskkkgalgpetsgce 67
QY 67 --WDGLICWPRTGKISAVPCPPVIYDFNHK-GVAFRHCNPNCTWDFMHSINKTNWYS 123
DB 68 glwdnmwscwpsapartvegcplllmslnkngslfrnctqdg-----wse-- 114

QY 124 DCLRFQPDISIG---KOEFCER-----LYVMYTVGYSISFGSLAVAILIIGYFRLH 173
DB 115 ---tfprpdlacgvninnnsfnerhayllkikvmtygysslamllvalsilcsfrlh 171

QY 174 CTRNVIHMLFVSEMLRATSFVKDRVVAHIGVKELESIMQDDPONSIBATSVDSQY 233
DB 172 ctrnyihmlfvfslralsnfikdav-----lfsdd-----vtycd-ahk 212

QY 234 IGCKTAVVMFYFLATNYWILVEGLYHNLIFVAFFSDTKYLMGFILIGWGFPAFVAA 293
DB 213 vgcikvmiffqycimanyawillveglylhtllaisffserkylgafvilgwgspafval 272

QY 294 WAVARATLADARCWELSA-GDIKWIYOAPILAAIGLNFILFNTVRVLATKIWETNAVGH 352
DB 273 waitrhflentgcwdinanawwvrgpvllslinliffilnrlmrkirtqetrgrs 332

QY 353 DTRKQYRKLAKTSLVLVFGVHYIVFVCLPHSFSTGLGWEIRMHCELPFNSFQGFVSII 412
DB 333 et-nhykrakstlilipfghiivfafsped-----amevqlffelalgfsqglvvavl 387

QY 413 YCYNGEVOAEVKMWSRWNLSDWKRTPPCGSRRCGSLVTIVTHTSSQSQ 464
DB 388 ycfingevqlvqkwrqwhiq-eiprlrpvafnnnsfnatngphtstkste 438

RESULT 15
ID R72506 standard; Protein; 458 AA.
AC R72506;
DT 29-NOV-1995 (first entry)
XX Porcine vasoactive intestinal peptide receptor protein.
DE Porcine vasoactive intestinal peptide receptor; pig; lung tissue; probe;
KW amplification; primer; PCR; vasodilation; gastrointestinal tract.
XX Sus scrofa.
XX EP648837-A.
XX 19-APR-1995.
XX 24-AUG-1994; 94EP-0306251.
XX 25-AUG-1993; 93US-0112817.
XX (ELIL) LILLY & CO ELI.
XX Hsiung HM, Smith DP, Zhang X;
XX WPI; 1995-148716/20.
XX N-PSDB; 089546.
XX Porcine vasoactive intestinal peptide receptor and DNA - useful in
PT receptor bioactivity assay and to screen for agents which inhibit or
PT stimulate receptor activity
XX

PS CC The amino acid sequence of the novel porcine vasoactive intestinal
XX peptide (pVIP) receptor. The gene was isolated from a cDNA library
CC in the Superscript (RTM) Lambda System derived from porcine lung tissue
CC RNA using a 700 bp probe amplified from lung tissue cDNA using primers
CC Q89547-8. The gene was inserted into the plasmid pSPC/CMV for
CC Thereafter the gene was inserted into the plasmid pSPC/CMV for
CC transfection of 293 cells and production of the recombinant protein.
CC The activity of pVIP is linked to vasodilation in the lungs and
CC gastrointestinal tract. The receptor is useful in a bioactivity assay
CC for quantifying the level of stimulation and repression of pVIP enzymatic
CC activity in response to test compounds. The receptor is also useful in
CC screening for pVIP receptor inhibitory or stimulatory agents.
XX Sequence 458 AA;
SQ

Query Match 27.0%; Score 785; DB 16; Length 458;
Best Local Similarity 37.0%; Pred. No. 1.5e-76;
Matches 182; Conservative 76; Mismatches 152; Indels 82; Gaps 18;

QY 12 GWMLGSCLLA---RAQLDSGDTIT---EQIVLVKAKVQCELNITAOQOEGEGNCF 64
DB 10 gwf-----cvlagvliacvlgpvgswavglqgeecdyllmaikvghkqcleeaqlenetsgcs 65

QY 65 PEWDGLICWPRTGKISAVPCPPVIYDFN-HKGV-AFRHCNPNCTWDFMHSINKTNWANY 122
DB 66 kmwdnltcwpatpgrgvvliacplflkfstptqglnvsnrctdeg-----wtp- 113

QY 123 SDCLRFLOP-----DISIGKQE---FCERLYVMYTVGYSISFGSLAVAILIIGY 168
DB 114 -----lepgypiacgmddkasgldeqctvfynsvktygtigyslslaallvatalisl 167

QY 169 FRRUHCTRNVIHMLFVSEMLRATSFVKDRVVAHIGVKELESIMQDDPONSIEATSV 228
DB 168 frklhctrnyihmlfislratavfikdlalf-----dseesdhc 209

QY 229 DKSQYIGCKIAVVMFYFLATNYWILVEGLYHNLIFVAFFSDTKYLMGFILIGWGFPA 288
DB 210 skgs-vgcgaavvlfqycvmanffwllvegylhlhllavsfserkyfwtgvyfvgwgvps 268

QY 289 AFVAAMAVARATLADARCWELSGADIKWIYOAPILAAIGLNFILFNTVRVLATKIWETN 348
DB 269 tfimvwtvtrihfedygcdtihsllwlikapilasilvnlflfirigilvqklrppd 328

QY 349 AVGHDTKQYRKLAKTSLVLVFGVHYIVFVCLPHSFSTGLGWEIRMHCELPFNSFQGF 408
DB 329 -vgksdnspysrlakstlilipfghvhyimfafpdnfka---evkmvfelivgsgqcv 384

QY 409 VSIIVCYNGEVOAEVKMWSRWNLSDWKRTPPCGSR---LSVDMKRTPPCGSR---CG---SVLTVT 456
DB 385 vallycfingevqaelrrkrwrwhqgvlgwdsqyhgpsgsgngdctcstqvmrlrvsps 444

QY 457 --HSTSSQSQVA 466
DB 445 arrsssfqaevs 456

Search completed: November 8, 2000, 08:51:37
Job time: 249 sec

XX PA (GEO) GEN HOSPITAL CORP OFFICE TECHNOLOGY.
 XX PI Abou-samra A, Juppner H, Kronenberg HM, Potts JT, Schipani E;
 XX PI Segre GV;
 XX DR WPI; 1992-366271/44.
 XX DR N-PSDB; Q29607.
 XX PT New DNA encoding parathyroid hormone receptor, DNA and antibodies
 XX PT - for (differential) diagnosis of hypercalcaemia, and diagnosis
 XX PT and treatment of tumours
 XX PS Claim 22; Fig 6; 91pp; English.
 XX CC The protein sequence was deduced from the cDNA sequence obtd by
 CC screening a human kidney oligo dt-primed cDNA library in lambda
 CC g10 and a genomic library of human placental DNA in EMBL3 with a
 CC probe comprising most of the coding sequence of rat bone parathyroid
 CC hormone/parathyroid hormone related protein (PTH/PTHrP) receptor
 CC protein. The clone encodes a protein which may be used in a
 CC therapeutic compsn. to inhibit activation of PTH or PTHrP and thus
 CC reduce the level of calcium in the blood. Cods. capable of competing
 CC with PTH or PTHrP for binding can be identified using the protein prod.
 CC and DNAs homologous to PTH DNA can be identified using fragments of the
 CC clone as probes. The sequence may be used for the prodn. of antibodies
 CC useful for the treatment, classification, prognosis and/or treatment of
 CC disorders related to the interaction between a cell receptor and a
 CC ligand such as in hypercalcaemia. See also R27704-16.
 XX SQ Sequence 614 AA;
 Query Match 44.8%; Score 1303.5; DB 13; Length 614;
 Best Local Similarity 50.2%; Pred. No. 9.9e-133;
 Matches 264; Conservative 68; Mismatches 117; Indels 77; Gaps 9;
 QY 22 ARAQLSDGRTTIEEQIVLVKAKVQCELNITAQLO----- 57
 Db 22 ayalvdaddvmtkeeqifllhraqaqcekrleqlvqpasimesdkgtwstsgkprkd 81
 QY 58 -----EGEGN-----CFPEWDGLICWPRGTGKISAVPCPPYIYDNHK 96
 Db 82 kasgklypseedkeaptgrygrpcpewdhilcplgapsevavpccdyldfnhk 141
 QY 97 GVAFRHCNPGTWDEHSLNKNWANYSDCLRFLOPDISIGKQFCERLYVYVTVGYSISF 156
 Db 142 ghayrrcdngtwelvpghnttwanyseckvftnetr--erevfdrlgmityvgsysl 199
 QY 157 GSLAVAILIIGYFRLRCHTRNYIHMHLFVSFMLRATSFYVKDRVYVHAHIGVKELES--- 213
 Db 200 asltvavllayfrlhrctnryihmhlfsfmlravsfvkdavpysgatldaeerltee 259
 QY 214 ----IMODDPQNSIATSVDSKQSYIGCKIAVWFYIFLATNYWILVEGLYLNHLIFVAF 269
 Db 260 elraiaqappataaag-----yagervavtfflylatnyywilvegilylnhlifmaf 314
 QY 270 FSDTKYLWGFTLIGWGFPAAFAVAWAVARATLADARCWELSGADIKWIOAPILAAIGLN 329
 Db 315 fsekkyllwgtfvfgwglpavfvavwsvratlantgdcwldssgnkklwllqvpilasvln 374
 QY 330 FILFLNTRVRLATKIWFTNAVGHDTKQYRKLAKTSLVLVLVFGVHVYIFVCLPFS--FTG 388
 Db 375 filfinivrlatkqretnagrcdrtqyckllkstivlmpflgvyivfmatpytevsg 434
 QY 389 LGWEIRHMCLEFNPSFGGFFVSIYCYNCGEVQAEVKKMWSNLSVDWRTPPCGRRRC 448
 Db 435 tlwqgmhyemlnsfsgffvliayfcfngvevqaeiikkswspwtlaidfkrrkarsg--- 491
 QY 449 GSVLTVTTHSTSSQSOVAHAHAWCLSLAK----LPRSPADSLTATS 490
 Db 492 -----stysyggpm-vstsvtnvgprglplsprrllpatt 527

RESULT 14
 R30187
 ID R30187 standard; Protein; 449 AA.
 XX AC R30187;
 XX DT 28-APR-1993 (first entry)
 XX DE Secretin receptor.
 XX KW Rat; rat/mouse hybridoma; NG 108-15.
 XX OS Rattus rattus.
 XX FH Location/Qualifiers
 FT Peptide 1..22 "signal peptide"
 FT Protein 23..449
 FT Modified-site 72 /note= "mature secretin receptor"
 FT Modified-site 100 /note= "potential N-glycosylation site"
 FT Modified-site 106 /note= "potential N-glycosylation site"
 FT Modified-site 128 /note= "potential N-glycosylation site"
 FT Modified-site 291 /note= "potential N-glycosylation site"
 FT Domain 144..165 /note= "potential N-glycosylation site"
 FT Domain 175..194 /note= "transmembrane domain"
 FT Domain 217..240 /note= "transmembrane domain"
 FT Domain 254..276 /note= "transmembrane domain"
 FT Domain 294..317 /note= "transmembrane domain"
 FT Domain 343..362 /note= "transmembrane domain"
 FT Domain 374..394 /note= "transmembrane domain"
 XX WO9221754-A.
 XX 10-DEC-1992.
 XX 05-JUN-1992; 92WO-JP00728.
 XX 07-JUN-1991; 91JP-0163946.
 XX (OSAB-) OSAKA BIOSCIENCE INST.
 XX Ishihara T, Nagata S, Takahashi K;
 XX WPI; 1992-433652/52.
 XX N-PSDB; Q33018.
 XX DNA coding for secretin receptor - is expressed in COS cells and
 XX produces a receptor protein for research and clinical use
 XX Claim 2; Fig 1; 44pp; Japanese.
 XX The secretin receptor was encoded by a DNA sequence of rat origin,
 XX contained in rat/mouse hybridoma NG108-15. The DNA sequence was
 XX obtd. from a cDNA library derived from NG108-15 cells. Expression
 XX in a suitable host allows prodn. of the receptor protein. The
 XX secretin receptor protein encoded by this gene may be used in basic
 XX research and in clinical tests, and is available in high yield.
 XX Sequence 449 AA;
 SQ

XX PA (GEOH) GEN HOSPITAL CORP.
 XX PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 XX PI Schipani E, Segre GV;
 XX XX
 DR WPI: 1999-034124/03.
 DR N-PSDB; V08391.
 XX XX
 PT Antibody to parathyroid hormone receptor - for diagnostic or
 PT therapeutic use
 PT
 XX Claim 7; Fig 6; 63pp; English.
 XX
 CC This sequence represents the human parathyroid hormone (PTH) receptor
 CC which is targeted by the antibody of the invention. The antibody
 CC of the invention is immunoreactive with naturally occurring human, rat or
 CC opossum PTH receptor. The antibody is useful for treating disorders
 CC characterised by overstimulation of PTH receptors by their ligand and for
 CC the diagnosis of PTH-related hypercalcaemia.
 XX
 SQ Sequence 593 AA;

Query Match 45.8%; Score 1331.5; DB 20; Length 593;
 Best Local Similarity 47.4%; Pred. No. 8.5e-136;
 Matches 276; Conservative 80; Mismatches 135; Indels 91; Gaps 11;

QY 22 ARAQLSDSGTITTEQIQLVLKAKVQCELNITAOLO----- 57
 DB 22 ayalvdaddvmtkeeqifllhraqaqcekrleqvlpasimesdkgwtastsqkprkd 81
 QY 58 -----EGEEN-----CFPEWDGLICWPGTGVKISAVPCPPPIYDFNHK 96
 DB 82 kasgklypeseeckaptgsrygrpcpewdhilcwpdgapgevavpcpdyidfnhk 141
 QY 97 GVAFRHCNPNGTWDFHSLNKTWANYSDCLRFLOPDISIGKQFCERLYVMYTVGSISF 156
 DB 142 ghayrrcdngswelvpghnrtwanyseckvftnetr--erevfdrlgmlycvgsyvl 199
 QY 157 GSLAVAILIIGYFRRLHCTRNYIHMHLFVSMFLRATSIYFKDRVVAHIGVKELES--- 213
 DB 200 asltvavllilayfrlhrctnryihmhlfsmlraasifvkdavlysgatldaeerlee 259
 QY 214 ----IMDDDPONSTEATSDVKSOYIGCKIAVVMFIYFLATNYWLLVEGLYHNLIFVAF 269
 DB 260 elraiaqapppataag-----yagcrvavtfflyfncnywlvveglyhnlifmaf 314
 QY 270 FSDTKYLMGFIIGWGPFAAFVAWAVARATLADARCWELSGADIKWIYQAPILAAIGLN 329
 DB 315 fsekylwgtfvfgwglpavfvavwsvratlantgcwdlssgnkwwliqvpilasivln 374
 QY 330 FILFLNTVRVLATKIWETNAVGHDTKQYRKLAkstlVLVLVFGVHVIVFVCLPHS-FTG 388
 DB 375 filflnivrvalatqretnagrcdtrqykrllkktvlImpflfgvhyivfmatpytevs 434
 QY 389 LGWEIRMHCELFNSFGFRVSIYCYVCNGEVOAEVKKWNRNLSVDWKRTPPCGSRR 448
 DB 435 tlwqgmnyemlfnsgqffvallyfcngcvgeaekkswrwtlaldfrkarssy 494
 QY 449 G-SVLTVTHSTSSQSOVAAHAWCLSLAKLPRSPADSLTATSL-YLAMSQVTSRT-AS 505
 DB 495 sygpmvshstvtngprvglg-----lpisprllptatngpqlghakpqtal 545
 QY 506 HTLSTRS-----NKEDSGRQRDILMEK 528
 DB 546 etletppamaapkdqfngscsgideeasgperpallqe 587

RESULT 11

R27706

ID R27706 standard; Protein: 591 AA.

XX

AC R27706;
 XX DT 16-MAR-1993 (first entry)
 XX DE Rat bone PTH/PTHrP receptor clone R15B prod.
 XX XX
 KW Parathyroid hormone; related protein; calcium; antagonist;
 KW antibodies; hypercalcaemia.
 XX OS Rattus rattus.
 XX XX
 PN WO9217602-A.
 XX PD 15-OCT-1992.
 XX XX
 PF 06-APR-1992; 92WO-US02821.
 XX PR 05-APR-1991; 91US-0681702.
 PR 06-APR-1992; 92US-0864475.
 XX XX
 PA (GEOH) GEN HOSPITAL CORP OFFICE TECHNOLOGY.
 XX XX
 PI Abou-samra A, Juppner H, Kronenberg HM, Potts JT, Schipani E;
 PI Segre GV;
 XX XX
 DR WPI: 1992-366271/44.
 DR N-PSDB; Q29606.
 XX XX
 PT New DNA encoding parathyroid hormone receptor, DNA and antibodies
 PT - for (differential) diagnosis of hypercalcaemia, and diagnosis
 PT and treatment of tumours
 XX
 PS Claim 20; Fig 3; 91pp; English.
 XX
 CC The rat bone parathyroid hormone/parathyroid hormone related
 CC protein (PTH/PTHrP) receptor protein sequence was deduced from
 CC clone R15B obtd. by screening a rat osteosarcoma (ROS) cell cDNA library
 CC to isolate those expressing functionally intact PTH/PTHrP receptor
 CC proteins, performed according to Gearing et al., (EMBO J. 8: 3676,
 CC 1989), by identifying colonies capable of binding a suitable radio-
 CC labelled ligand. The protein may be used in a therapeutic compsn. to
 CC inhibit activation of PTH or PTHrP and thus reduce the level of calcium
 CC in the blood. Cpd. capable of competing with PTH or PTHrP for binding
 CC can be identified using the protein and DNAs homologous to PTH DNA can
 CC be identified using fragments of the clone as probes. The sequence
 CC may be used for the prodn. of antibodies useful for the treatment,
 CC classification, prognosis and/or treatment of disorders related to
 CC the interaction between a cell receptor and a ligand such as in
 CC hypercalcaemia. See also R27704-16.
 XX
 SQ Sequence 591 AA;

Query Match 45.5%; Score 1323.5; DB 13; Length 591;
 Best Local Similarity 46.5%; Pred. No. 6.2e-135;
 Matches 269; Conservative 86; Mismatches 137; Indels 87; Gaps 11;
 QY 22 ARAQLSDSGTITTEQIQLVLKAKVQCE-----LNITAOLOEGEEN----- 62
 DB 22 ayalvdaddvmtkeeqifllhraqaqcdkllkevlhtaanimesdkgwtastsgkprke 81
 QY 63 -----CFPEWDGLICWPGTGVKISAVPCPPPIYDFNHK 96
 DB 82 kasgklypesekenkdvptgsrrgrpcpewdnivcwpdgapgevavpcpdyidfnhk 141
 QY 97 GVAFRHCNPNGTWDFHSLNKTWANYSDCLRFLOPDISIGKQFCERLYVMYTVGSISF 156
 DB 142 ghayrrcdngswelvpghnrtwanysecklkmtnetr--erevfdrlgmlytvgsyml 199
 QY 157 GSLAVAILIIGYFRRLHCTRNYIHMHLFVSMFLRATSIYFKDRVVAHIGVKELES--- 213
 DB 200 asltvavllilayfrlhrctnryihmhlfsmlraasifvkdavlysgatldaeerlee 259

```
Qy 157 GSLAVAILIIGYFRRLHCTRNYYIHMHVFSVPMRLRATSFVKDRVVHAHIGVKELES--- 213
Db 200 asltvavllayfrlhcrtynyihmhflsfmlraasfvkdavlysgftldeaeerlee 259
Qy 214 ----IMQDDPQNSIEATSDKSOYIGCKIAVVMFIYFLATNYWILVEGLYLNLIIFVAF 269
Db 260 elhiiaqvpppppaaavg----yagervavtfflyflatnywilylveglylhlifmaf 314
Qy 270 FSDTKYLWGFLLIGWFPAAFAAWAVARATLADARCWELSGADIKWYQAPIAAIAGLN 329
Db 315 fsekkylgwftifgwpavfvavvgratlangtgcwdisgghkwiqvpllasvvl 374
Qy 330 FILELNTVRVLATKIWETNAVGHDRKQYRKLAKESTLVLVLFVGVHYIVFVCLPHS-FTG 388
Db 375 filfiniirvatklretnagrcdrqyrklrstlilvplfgvhytvfmalpytevsg 434
Qy 389 LGWEIRHCELFNFSGFQFFVSIYCYCNGEVOAEVKMMSRWNLSDWKRTPPCGSRR 448
Db 435 tlwqimhyemlfnsgffvatiycfngvqaeirkswrtlaidfrkarsgssy 494
Qy 449 GSVLTVTHTSTSSQVAAAHAMCLSLAKLPRSPADSLTATSLYLAMSG-----VTQS 501
Db 495 -sygpmvshstvtngvrag----lsiplsprlpp---attnghsqipghakpgapatet 546
Qy 502 RTASHTLSTRS-----NKEDSGRQRDDILMEK 528
Db 547 etlptmavpkddgflngscsgldeasgsarpppliqe 585

RESULT 9
W73316 ID W73316 standard; Protein; 591 AA.
XX
AC W73316;
XX
DT 08-FEB-1999 (first entry)
XX
DE Parathyroid hormone receptor R15B.
XX
KW Parathyroid hormone receptor; PTH receptor; antibody; therapy;
KW PTH-related hypercalcaemia; rat.
XX
OS Rattus sp.
XX
PN US5840853-A.
XX
PD 24-NOV-1998.
XX
PF 06-JUN-1995; 95US-0471494.
XX
PR 06-APR-1992; 92US-0864475.
PR 05-APR-1991; 91US-0681702.
PR 06-JUN-1995; 95US-0471494.
XX
PA (GHEO ) GEN HOSPITAL CORP.
XX
PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
PI Schipani E, Segre GV;
XX
XX WPI; 1999-034124/03.
DR N-PSDB; V08390.
XX
XX Antibody to parathyroid hormone receptor - for diagnostic or
PT therapeutic use
XX
PS Claim 6; Fig 3; 63pp; English.
XX
CC This sequence represents the rat parathyroid hormone (PTH) receptor
CC R15B, which is targeted by the antibody of the invention. The antibody
CC of the invention is immunoreactive with naturally occurring human, rat
CC opussum PTH receptor. The antibody is useful for treating disorders
CC characterised by overstimulation of PTH receptors by their ligand and for
CC the diagnosis of PTH-related hypercalcaemia.
```

```
XX
SQ Sequence 591 AA;
Query Match 45.8%; Score 1331.5; DB 20; Length 591;
Best Local Similarity 46.6%; Pred. No. 8.4e-136;
Matches 270; Conservative 86; Mismatches 136; Indels 87; Gaps 11;
Qy 22 ARAOLDSDGRTTIEQIVLVLKAKVQCE-----LNITAOLQEGEGN----- 62
Db 22 ayalvdaddvftkeeqflhraqacdkllkevlhtaanlmesdkgwtpastsgkprke 81
Qy 63 -----CFPEWDGLICWPRGTGVKISAVPCPPYIYDFNMH 96
Db 82 kasgkfypeskenkdvtgsrrirgpcelpewdnivcplgaapgevavpcpdyldfnhk 141
Qy 97 GVAPRHCPNCTWDFMISLKNKTWANYSDCLFLQDPDISIGKQFCERLYVMWYTVGYSISF 156
Db 142 ghayrrcdrgswvvpghnrtwanyscclkmfnetr--erevfdrlgmiytvgysmsl 199
Qy 157 GSLAVAILIIGYFRRLHCTRNYYIHMHVFSVPMRLRATSFVKDRVVHAHIGVKELES--- 213
Db 200 asltvavllayfrlhcrtynyihmhflsfmlraasfvkdavlysgftldeaeerlee 259
Qy 214 ----IMQDDPQNSIEATSDKSOYIGCKIAVVMFIYFLATNYWILVEGLYLNLIIFVAF 269
Db 260 elhiiaqvpppppaaavg----yagervavtfflyflatnywilylveglylhlifmaf 314
Qy 270 FSDTKYLWGFLLIGWFPAAFAAWAVARATLADARCWELSGADIKWYQAPIAAIAGLN 329
Db 315 fsekkylgwftifgwpavfvavvgratlangtgcwdisgghkwiqvpllasvvl 374
Qy 330 FILELNTVRVLATKIWETNAVGHDRKQYRKLAKESTLVLVLFVGVHYIVFVCLPHS-FTG 388
Db 375 filfiniirvatklretnagrcdrqyrklrstlilvplfgvhytvfmalpytevsg 434
Qy 389 LGWEIRHCELFNFSGFQFFVSIYCYCNGEVOAEVKMMSRWNLSDWKRTPPCGSRR 448
Db 435 tlwqimhyemlfnsgffvatiycfngvqaeirkswrtlaidfrkarsgssy 494
Qy 449 GSVLTVTHTSTSSQVAAAHAMCLSLAKLPRSPADSLTATSLYLAMSG-----VTQS 501
Db 495 -sygpmvshstvtngvrag----lsiplsprlpp---attnghsqipghakpgapatet 546
Qy 502 RTASHTLSTRS-----NKEDSGRQRDDILMEK 528
Db 547 etlptmavpkddgflngscsgldeasgsarpppliqe 585

RESULT 10
W73317 ID W73317 standard; Protein; 593 AA.
XX
AC W73317;
XX
DT 08-FEB-1999 (first entry)
XX
DE Human Parathyroid hormone receptor.
XX
KW Parathyroid hormone receptor; PTH receptor; antibody; therapy;
KW PTH-related hypercalcaemia; human.
XX
OS Homo sapiens.
XX
PN US5840853-A.
XX
PD 24-NOV-1998.
XX
PF 06-JUN-1995; 95US-0471494.
XX
PR 06-APR-1992; 92US-0864475.
PR 05-APR-1991; 91US-0681702.
PR 06-JUN-1995; 95US-0471494.
```



```

Qy 342 TKIWTNAVGHDTKQYRKRLAKSTLVLVVFGVHYIVFVCLPHS-FTGLGWEIRMHCELF 400
Db 381 tklnetnagrcdtrqykrkllkstlvimplfgvhyivfmatpytevsgilwqgmhyeml 440
Qy 401 FNSFGGFVSIYCYCNGEVOAEVKKMSRWNLSDWKRTPPCGSRRCGSLTIVTTH--- 457
Db 441 instqgfvaliycfcngevqaeikksrwltalldfkrkarsgs-----stysygm 493
Qy 458 -STSSQSOVAAAHAWCLSLAKLPRSPADSLTATSLYLAMSGVTSRTASHTLSRKNED 516
Db 494 vshtsvtnvgrgglalsl--prlapgagasanghhqlpgyvkhsisen-slpssgpe 550
Qy 517 SGQRDDILM-----EKPSRPMESNPD 539
Db 551 pgtkddgylngsglyepmvgeqpplleereet 583

RESULT 3
W73315
ID W73315 standard; Protein; 585 AA.
AC W73315;
DT 08-FEB-1999 (first entry)
DE Parathyroid hormone receptor OK-O.
KW Parathyroid hormone receptor; PTH receptor; antibody; therapy;
KW PTH-related hypercalcaemia; opossum.
XX OS
XX OS
XX Didelphis virginiana.
XX US5840853-A.
XX 24-NOV-1998.
XX 06-JUN-1995; 95US-0471494.
XX 06-APR-1992; 92US-0864475.
XX 05-APR-1991; 91US-0681702.
XX 06-JUN-1995; 95US-0471494.
XX (GEO ) GEN HOSPITAL CORP.
XX Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
XX Schipani E, Segre GV;
XX WPI; 1999-034124/03.
XX N-PSDB; V08389.
XX Antibody to parathyroid hormone receptor - for diagnostic or
XX therapeutic use
XX Claim 6; Fig 2; 63pp; English.
XX This sequence represents the opossum parathyroid hormone (PTH) receptor
XX OK-O, which is targeted by the antibody of the invention. The antibody
XX of the invention is immunoreactive with naturally occurring human, rat or
XX opossum PTH receptor. The antibody is useful for treating disorders
XX characterised by overstimulation of PTH receptors by their ligand and for
XX the diagnosis of PTH-related hypercalcaemia.
XX Sequence 585 AA;
SQ
Query Match 47.9%; Score 1392; DB 20; Length 585;
Best Local Similarity 47.8%; Pred. No. 2.2e-142;
Matches 274; Conservative 92; Mismatches 137; Indels 70; Gaps 9;

Qy 24 AQLSDGTITIEQIVLVKAKVOCELMITAQLQEGE----- 60
Db 24 alvdaddvitkeeqillirnaqagceqrkveirvpelaesakdmsrsaktkekpaek 83

```

```

Qy 61 -----GNCFFPEWGLICWPRGTGKISAVPCPPYIYDFNHNKGVAFR 101
Db 84 lypaaeesrevsdrslqdgfcelpewdnivcwpagvpkgvavpcpdyfydfnhkgrayr 143
Qy 102 HCNPNGTWDFMHSLNKTWANYSDCLRFLOPDISIGKQFCEERLYVMYTVGVYSISFGSLAV 161
Db 144 rcdngswelypgnrtrtwanyseckfltnetr--erevfdrlgmlytvgyvislgsitv 201
Qy 162 AILLIGYFRRLHCTRNYYTHMLFVSEMLRATSIFFKORVVHAHIGVKKELESIMQDDPQN 221
Db 202 avlligyfrfllhctrnymhmlfsvfmlravsfikdaviysgvstdeier-iteeeira 260
Qy 222 SIETATSDKSOYIGCKIAVVMFIYFLATNYXWILVEGLYHNLIFVAFPSDTKYLWGFIL 281
Db 261 fteppackagfvcgrvavtvflyfltnyywilvegylhslifmaffsekkylwgftl 320
Qy 282 IGMGFPAFVAANAVARATLADARCWELSGADIKWIYQAPILAAIGLNFILFNTVRVLA 341
Db 321 fgwlpavfvavvtvratlantecwdlssgnkklwllqvpilaaivvnflfinlirvla 380
Qy 342 TKIWTNAVGHDTKQYRKRLAKSTLVLVVFGVHYIVFVCLPHS-FTGLGWEIRMHCELF 400
Db 381 tklnetnagrcdtrqykrkllkstlvimplfgvhyivfmatpytevsgilwqgmhyeml 440
Qy 401 FNSFGGFVSIYCYCNGEVOAEVKKMSRWNLSDWKRTPPCGSRRCGSLTIVTTH--- 457
Db 441 fnsqgfvaliycfcngevqaeikksrwltalldfkrkarsgs-----stysygm 493
Qy 458 -STSSQSOVAAAHAWCLSLAKLPRSPADSLTATSLYLAMSGVTSRTASHTLSRKNED 516
Db 494 vshtsvtnvgrgglalsl--prlapgagasanghhqlpgyvkhsisen-slpssgpe 550
Qy 517 SGQRDDILM-----EKPSRPMESNPD 539
Db 551 pgtkddgylngsglyepmvgeqpplleereet 583

RESULT 4
R27705
ID R27705 standard; Protein; 585 AA.
XX AC
XX R27705;
XX 16-MAR-1993 (first entry)
XX Opossum kidney PTH/PTHrP receptor prod. from clone OK-O.
XX Parathyroid hormone; related protein; calcium; antagonist;
XX antibodies; hypercalcaemia.
XX Didelphis virginiana.
XX W09217602-A.
XX 15-OCT-1992.
XX 06-APR-1992; 92MO-US02821.
XX 05-APR-1991; 91US-0681702.
XX 06-APR-1992; 92US-0864475.
XX (GEO ) GEN HOSPITAL CORP OFFICE TECHNOLOGY.
XX Abou-samra A, Juppner H, Kronenberg HM, Potts JT, Schipani E;
XX Segre GV;
XX WPI; 1992-366271/44.
XX N-PSDB; Q29605.
XX New DNA encoding parathyroid hormone receptor, DNA and antibodies
XX - for (differential) diagnosis of hypercalcaemia, and diagnosis
XX and treatment of tumours
PT

```

PT hyper-calcaemia, hypo- or hyper-phosphatemia, kidney stones, etc
 XX
 PS Claim 9; Fig 1A-E; 62pp; English.
 XX
 CC A novel 7-transmembrane receptor (W12695) has been identified as a
 CC human G-protein parathyroid hormone (PTH) receptor, designated
 CC HLTG74. It shows 48.2% homology to the human PTH receptor. Its
 CC amino acid sequence was deduced from a cDNA clone (T59619) isolated
 CC from a human T cell lymphoma tissue cDNA library. Recombinant
 CC HLTG74 can be produced in transformed host cells and used to
 CC screen for (ant)agonist cpds. Agonists may be used to prevent or
 CC treat e.g. hypocalcaemia, hyperphosphataemia, hypoparathyroidism
 CC and chronic tetany by stimulating an increase in serum calcium
 CC levels. Antagonists can be used to inhibit the receptor e.g. for
 CC the treatment of osteoporosis, hypercalcaemia, hypoparathyroidism,
 CC hypophosphataemia, kidney stone, nephrolithiasis.
 XX
 SQ Sequence 541 AA;

Query Match 100.0%; Score 2907; DB 18; Length 541;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAWLGSLHWGVLMLGSCLLARQLSDGTITTEEQIVLVKAKVOCENITQAQLEGE 60
 DB 1 mawlgaslhvhwgvlmlgscllaraqlsdgtitteeqivlvkavqcenitlaqlqege 60
 QY 61 GNCPEPDGGLICWPRGTGKISAVPCPPYLYDFNHKGVAFRHCNPNCTWDFMHSLNKTWA 120
 DB 61 gncfpewdgglcwpgrgtgkisavpcppylydfnkhgvafrhcnpngctwdfmhslnkntwa 120
 QY 121 NYSCLRFLOPDISIGKOEFCERLYVMYTVGYISFGSLAVAILIIGYFRRLHCTRYIH 180
 DB 121 nysclrlflopdisigkoeerlyvmtyvgyisfgslavailiigyfrlhrctryih 180
 QY 181 MHLFVSPMLRATSFVVDKRVVHAHIGVKELSLIMQDDPONSIEATSVDSQYIGCKIAV 240
 DB 181 mhlfsvmlratslfvvdkrvvhahigvkeleslimqddpqnslieatsvdsqyigckia 240
 QY 241 VMFYFLATNYWTLVGLYLHNLIFVAFSDTKYLMGFILIGWGPAAFAVAWAVARAT 300
 DB 241 vmfyflatnywtilvgylylhnlfvafsdtkylmgfildgwpaaafaawavarat 300
 QY 301 LADARCWELSGADIKTIYQAPILAAIGLNFILFLNTVRLATKIWETNAVGHDTKQYRK 360
 DB 301 ladarcwelsagdiktyqapilaaiglnfllntvrvlatkiwetnavghdtkqyrk 360
 QY 361 LAKSTLVVLVFGVHYTVFVCLPHSFTGLGWETRMHCELFNSFGQFFVSLIYCYCNGEV 420
 DB 361 lakstlvvlvfgvhytvfvcplhpsftglgwetrmhcelfnsgffvsllycycngev 420
 QY 421 QAEVKKMWRSWNLSDVMKTRPPCGSRRCSGLTTLVTHSTSSQVAAAHAWCLSLAKLPR 480
 DB 421 qaevkkmwrswnlsdvmktrppcgssrrcsglvtlvtthstssqvaahawclslaklpr 480
 QY 481 SPADSLTATSLYLAMSGVTSQTSASHTLSTRSNKEDSGRQDDTLMEKPSRPMESNPDE 540
 DB 481 spadsltatslylamsgvtsqtsashtlstrsnkedsgrqddtlmekpsrpmesnpdte 540
 QY 541 G 541
 DB 541 g 541

RESULT 2
 R92276
 ID R92276 standard; Protein: 585 AA.
 XX
 AC R92276;
 XX
 DT 18-MAY-1996 (first entry)
 XX

DE Opossum kidney PTH/PTHrP receptor.
 XX
 KW Parathyroid hormone; receptor; parathormone; PTH;
 KW parathyroid hormone-related protein; PTHrP; calcium; homeostasis;
 KW hypercalcaemia; hypocalcaemia; cancer; opossum.
 XX
 OS Didelphis virginiana.
 XX
 PN US5494806-A.
 XX
 PD 27-FEB-1996.
 XX
 PF 05-APR-1991; 91US-0681702.
 XX
 PR 06-APR-1992; 92US-0864475.
 PR 05-APR-1991; 91US-0681702.
 XX
 PA (GEO) GEN HOSPITAL CORP.
 XX
 PI Abou-Samra A, Juppner H, Kronenberg HM, Potts JT;
 PI Schipani E, Segre GV;
 XX
 DR WPI: 1996-139028/14.
 DR N-PSDB; T15946.
 XX
 PT DNA encoding vertebrate parathyroid hormone receptor - useful for
 PT diagnosis and treatment of e.g. hypercalcaemia, hypo-calcaemia,
 PT cancer etc.
 XX
 PS Claim 1; Fig 2A-2E; 64pp; English.
 XX
 CC Parathyroid hormone/parathyroid hormone-related protein (PTH/PTHrP)
 CC receptors (R92275 and R92776) are encoded by cDNA clones OK-H (T15945)
 CC and OK-O (T15946), respectively, isolated from an opossum kidney (OK)
 CC cell cDNA library. OK-H and OK-O may be the products of 2 separate
 CC genes or of a laboratory artifact. The receptor induces an increase
 CC in intracellular cAMP and calcium when challenged with PTH or PTHrP.
 CC Recombinant receptors can be produced in vector/host cell systems and
 CC used in the treatment, diagnosis or prognosis of e.g. hypercalcaemia
 CC and hypocalcaemia, to screen for (ant)agonists and to raise antibodies.
 CC Host cells expressing the receptor can be used for diagnostic
 CC measurement of PTH serum levels.
 XX
 SQ Sequence 585 AA;

Query Match 47.9%; Score 1392; DB 17; Length 585;
 Best Local Similarity 47.8%; Pred. No. 2.2e-142;
 Matches 274; Conservative 92; Mismatches 137; Indels 70; Gaps 9;

QY 24 AQLSDSGTITIEQIVLVKAKVOCENITQAQLEGE----- 60
 DB 24 alvdaddvikeeqiillrnaqqceqrilevirvpelaesakdmsrskttkekpae 83
 QY 61 -----GNCFFPNDGLICWPRGTGKISAVPCPPYLYDFNHKGVAFR 101
 DB 84 lypqaeesrevsdrsrldqgfcipewdnivcwpagvpgkvavpcpdyfydfnhkgrayr 143
 QY 102 HCNPNGTWDFMHSLNKTWANYSDCLRFLOPDISIGKOEFCERLYVMYTVGYISFGSLAV 161
 DB 144 rcdsngswelvpgnrtwanysecvklftnetr--erevdfirmiytvgyislsaltv 201
 QY 162 AILIIGYFRBLHCTRYIHMHFLVSPMLRATSFVVDKRVVHAHIGVKELSLIMQDDPON 221
 DB 202 aviligyfrlhrctryihmhflvsvfmlravsfikdavlsgvstdeier-lceela 260
 QY 222 SIEATSVDSKSOYIGCKTAVVMFYFLATNYWTLVGLYLHNLIFVAFSDTKYLMGFIL 281
 DB 261 fteppadkagfvcgrvavtvflylfttnyywllveglylhlslfmaffsekylwqf 320
 QY 282 IGWGFPAAFVAANAVARATLADARCWELSGADIKTIYQAPILAAIGLNFILFLNTVRLA 341
 DB 321 fgwglpavfvavvtvratlantecwldssgnkwiqvpilaavvfnlfnlirvla 380

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2000, 08:47:28 ; Search time 138.73 Seconds
(without alignments)
133.344 Million cell updates/sec

Title: US-09-236-468a-2

Perfect score: 2907

Sequence: 1 MAWLGLASLHWGWLMLGSL.....DDILMEKSRPMESNPDTREG 541

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_36.*
- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.*
 - 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.*
 - 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.*
 - 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.*
 - 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.*
 - 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.*
 - 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.*
 - 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.*
 - 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.*
 - 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.*
 - 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.*
 - 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.*
 - 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.*
 - 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.*
 - 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.*
 - 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.*
 - 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.*
 - 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.*
 - 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.*
 - 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.*
 - 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2907	100.0	541	18	W12695
2	1392	47.9	585	17	R92276
3	1392	47.9	585	20	W73315
4	1391	47.9	585	13	R27705
5	1388	47.7	515	17	R92275
6	1388	47.7	515	20	W73314
7	1375	47.3	515	13	R27704
8	1331.5	45.8	591	17	R92277
9	1331.5	45.8	591	20	W73316
10	1331.5	45.8	593	20	W73317
11	1323.5	45.5	591	13	R27706
12	1313	45.2	593	17	R92278

13	1303.5	44.8	614	13	R27707
14	801	27.6	449	13	R30187
15	785	27.0	458	16	R72506
16	772.5	26.6	459	14	R42848
17	713.5	24.5	437	19	W80309
18	713.5	24.5	437	20	W92973
19	708	24.4	437	16	R70136
20	704	24.2	431	19	W80310
21	704	24.2	431	20	W92974
22	702.5	24.2	438	19	W80308
23	702.5	24.2	438	20	W92972
24	696.5	24.0	448	15	R58666
25	696.5	24.0	467	15	R58656
26	689.5	23.7	525	15	R58659
27	688.5	23.7	448	15	R58665
28	688.5	23.7	485	15	R58657
29	684.5	23.5	448	15	R58668
30	681.5	23.4	476	15	R58667
31	681.5	23.4	495	15	R58658
32	675	23.2	552	15	R58661
33	674.5	23.2	553	15	R58660
34	673.5	23.2	476	15	R58663
35	673.5	23.2	513	15	R58655
36	672.5	23.1	553	15	R58662
37	670	23.0	475	15	R58670
38	669.5	23.0	476	15	R58669
39	667.5	23.0	476	15	R58671
40	665	22.9	550	19	W68066
41	655	22.5	463	14	R41877
42	640.5	22.0	466	21	Y51526
43	635.5	21.9	553	19	W68065
44	632	21.7	445	21	Y82704
45	631	21.7	418	15	R51071

ALIGNMENTS

RESULT 1

W12695

ID W12695 standard; Protein; 541 AA.

XX

AC W12695;

XX

DT 31-MAY-1997 (first entry)

XX

DE G-protein parathyroid hormone receptor HLTGD74.

XX

KW G-protein parathyroid hormone receptor; HLTGD74; parathormone; PTH;

KW calcium; signal transduction; agonist; antagonist; hypocalcaemia;

KW hyperphosphataemia; hypoparathyroidism; chronic tetany;

KW osteoporosis; hypercalcaemia; hypoparathyroidism; hypophosphataemia;

KW kidney stone; nephrolithiasis; therapy; diagnosis.

XX

OS Homo sapiens.

XX

PN WO9639433-A1.

XX

PD 12-DEC-1996.

XX

PF 05-JUN-1995; 95WO-US07085.

XX

PR 05-JUN-1995; 95WO-US07085.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Li Y, Rosen CA, Ruben SM, Soppet DR;

XX

WPI; 1997-043068/04.

DR

DR N-PSDB; T59619.

XX

PT Human G-protein parathyroid hormone receptor, HLTGD74 - used to identify (ant)agonists, used in the treatment of hypo- or